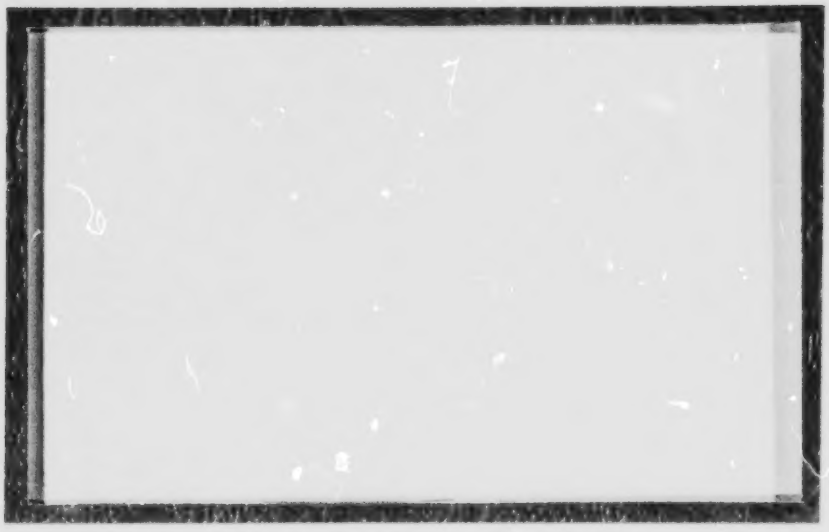
UNIVERSITÉ DE MONTRÉAL



CENTRE DE RECHERCHE SUR LES TRANSPORTS

MÉMOIRE D'ÉTUDIANT

CASE POSTALE 6128, SUCCURSALE A, MONTRÉAL, QUÉBEC, H3C 3J7

LA DEMANDE D'ESSENCE AU QUEBEC: UNE ETUDE DES FORMES FONCTIONNELLES

par
Lina Di Pietro

Ce travail a été subventionné par le Centre de recherche et de développement de Transports Canada et par le programme F.C.A.C. du Ministère de l'éducation du Québec.

Ce rapport a été publié grâce à une subvention du fonds F.C.A.C. pour l'aide et le soutien à la recherche.

Université de Montréal
Centre de recherche sur les transports - Publication #154
Août 1979

Dépôt légal - 1er trimestre 1980

UNIVERSITÉ DE MONTRÉAL

RÉSUMÉ

Le présent rapport de recherche a pour objectif l'étude de la détermination conjointe de la forme fonctionnelle et de la spécification de l'erreur pour une fonction économique spécifique. Les transformations paramétriques Box-Cox à puissance directe et inverse appliquées à la partie stochastique du modèle permettent de déterminer une forme fonctionnelle de l'hétéroscédasticité des résidus extrêmement générale (Gaudry et Dagenais, 1978). Cette forme fonctionnelle qui caractérise l'hétéroscédasticité est étudiée avec un modèle dont la partie fixe est elle-même sujette à des transformations directes.

Le modèle présenté utilise les séries chronologiques et de ce fait, la procédure Box-Cox appliquée aux parties fixe et stochastique du modèle est généralisée à une structure d'autocorrélation des résidus d'ordre multiple. L'échantillon utilisé pour notre fonction de demande d'essence dans la province de Québec est constitué de données mensuelles pour la période de décembre 1956 à juin 1977.

Nos résultats montrent que la détermination simultanée de la forme fonctionnelle optimale de la partie fixe et de la partie aléatoire a des effets importants sur les estimations des paramètres du modèle considéré.

TABLE DES MATIERES

RESUME i

INTRODUCTION 1

CHAPITRE I : Etude de la forme fonctionnelle de la partie fixe et de la spécification de l'erreur: revue de la littérature 3

1.1. Cas classique 4

1.1.1. Analyse de l'hétéroscédasticité du terme d'erreur . . . 5

1.1.2. Analyse de l'autocorrélation du terme d'erreur 7

1.2. Analyse de la forme fonctionnelle de la partie fixe du modèle (avec $E(uu') = \psi^2 I$) 8

. Transformation Box-Cox 9

1.3. Analyse de la forme fonctionnelle de la partie fixe simultanément avec la partie stochastique du modèle (avec $E(uu') \neq \psi^2 I$) 11

1.3.1. Forme de l'hétéroscédasticité dans un modèle linéaire 13

1.3.1.1. Une forme générale pour les erreurs hétéroscédastiques 13

1.3.1.2. La fonction de vraisemblance pour un modèle linéaire 16

1.3.1.3. La fonction de vraisemblance pour un modèle non-linéaire 17

1.3.2. Formulation complète d'un modèle non-linéaire en présence de l'hétéroscédasticité et d'une structure d'autocorrélation 19

1.3.2.1. du premier ordre 19

1.3.2.1.1. La fonction de vraisemblance en présence de l'hétéroscédasticité et de l'autocorrélation du premier ordre 20

1.3.2.1.2. Une formulation simplifiée de la vraisemblance en présence de l'autocorrélation du premier ordre . . . 22

1.3.2.1.3. Une formulation simplifiée de la vraisemblance en présence d'une structure d'autocorrélation multiple 24

CHAPITRE II : Procédure d'estimation 26

CHAPITRE III : Modèle à estimer 32

3.1. Application du modèle général à une fonction de demande d'essence 32

3.1.1. Paramétrisation de la partie fixe du modèle 33

3.1.2. Paramétrisation de la partie aléatoire du modèle 36

3.2. Nomenclature des hypothèses 37

3.2.1. Hypothèses sur la partie fixe 38

3.2.2. Hypothèses sur la partie aléatoire 39

3.3. Comparaison entre les hypothèses 43

. absence d'autocorrélation 43
. présence d'autocorrélation 44

CHAPITRE IV : Analyse des résultats 49

4.1. Etude sur le degré d'opérationnalité du modèle 49

4.2. Estimation et tests de la forme fonctionnelle 52

4.2.1. Justification de l'utilité de la détermination conjointe, à partir d'une analyse des différents modèles sur la base des hypothèses émises sur la partie stochastique du modèle 52

a) Deux variables prises séparément 53
b) Deux variables prises conjointement 61

4.2.2. Utilité de la généralité de la paramétrisation, à partir d'une analyse des différents modèles sur la base des hypothèses émises sur la partie fixe du modèle 63

4.2.3. Détermination du modèle qui maximise $L(\omega^{**})$ 68

4.3. Effets de la forme fonctionnelle et de la spécification de l'erreur 70

4.3.1. Effets de la forme fonctionnelle et de la spécification de l'erreur sur la distribution des résidus autorégressifs 70

. Coefficient de corrélation multiple 73

4.3.2. Effets de la forme fonctionnelle et de la spécification de l'erreur sur la signification des variables 73

4.3.3. Effets de la forme fonctionnelle et de la spécification de l'erreur sur les élasticités 84

CONCLUSION 95

APPENDICE 1 : Logarithme de la fonction de vraisemblance concentrée . . 97

APPENDICE 2 : Hypothèse de normalité de l'erreur 101

APPENDICE 3 : 3 mesures de coefficients de corrélation multiple 105

APPENDICE 4 : Définition des variables 107

BIBLIOGRAPHIE 109

LISTE DES TABLEAUX

TABLEAU 1 : Nombre d'observations positives pour chacun des régresseurs 35

TABLEAU 2 : Tableau récapitulatif sur les combinaisons d'hypothèses sur les parties fixe et aléatoire 41

TABLEAU 3 : Nombre d'itérations, temps et coût de résolution de divers modèles 51

TABLEAU 4 : Valeurs des paramètres λ_y, λ_{xk}, ρ_ℓ, λ_{em}, δ_m, avec $X_m \equiv$ *RSIDCB* et valeurs correspondantes des fonctions de vraisemblance. - Séries * et A - 54

TABLEAU 5 : Valeurs des paramètres λ_y, λ_{xk}, ρ_ℓ, λ_{em}, δ_m, avec $X_m \equiv$ *RSIDCB* et valeurs correspondantes des fonctions de vraisemblance. - Séries * et A - 57

TABLEAU 6 : Valeurs des paramètres λ_y, λ_{xk}, ρ_ℓ, λ_{em}, δ_m, avec $X_m \equiv$ *EIQP* et valeurs correspondantes des fonctions de vraisemblance. - Séries * et A - 60

TABLEAU 7 : Valeurs des paramètres λ_y, λ_{xk}, ρ_ℓ, λ_{em}, δ_m, avec $X_m \equiv$ *RSIDCB*, *EIQP* et valeurs correspondantes des fonctions de vraisemblance. - Séries A et AA - 62

TABLEAU 8 : Valeurs des paramètres λ_y, λ_{xk}, ρ_ℓ, λ_{em}, δ_m et valeurs correspondantes des fonctions de vraisemblance. Modèle BC_SAU - Séries *, A, B, C, D, E. - 64

TABLEAU 9 : Valeurs des paramètres λ_y, λ_{xk}, ρ_ℓ, λ_{em}, δ_m et valeurs correspondantes des fonctions de vraisemblance. Modèle BC_SAUHG_1 - Séries *, A, B, C, D, E. - 65

TABLEAU 10 : Valeurs des tests statistiques du rapport du maximum de vraisemblance pour les paramètres λ_y, λ_{xk}. Modèle BC_SAU - Séries *, A, B, C, D, E. - 66

TABLEAU 11 : Valeurs des tests statistiques du rapport du maximum de vraisemblance pour les paramètres λ_y, λ_{xk}. Modèle BC_SAUHG_1 - Séries *, A, B, C, D, E. - 66

TABLEAU 12 : Valeurs des paramètres λ_y, λ_{xk}, ρ_ℓ, λ_{em}, δ_m du modèle qui optimise $L(\omega^{**})$ 69

TABLEAU 13 : Fonctions d'autocorrélation et d'autocorrélation partielle, en différence des résidus d'ordre 0 . . . . . 72A

TABLEAU 14 : Valeurs des coefficients paramétriques, des élasticités, des tests statistiques et de $L(\omega^{**})$ pour divers modèles 74A

TABLEAU 15 : Nombre de paires de coefficients de corrélation partielle pour différents niveaux α 83

TABLEAU 16 : Rapport des élasticités du temps d'attente sur le temps de transport pour quelques modèles (en valeur absolue) 90

TABLEAU 17 : Rapport des élasticités du temps de trajet en automobile sur le prix de l'essence pour quelques modèles (en valeur absolue) 91

TABLEAU 18 : Rapport des élasticités du temps d'attente sur le temps de trajet en automobile pour quelques modèles (en valeur absolue) 91

TABLEAU 19 : Valeur implicite des prix du temps pour quelques modèles ($/heure) - p ≡ PGRMD - 92

TABLEAU 20 : Valeur implicite des prix du temps pour quelques modèles ($/heure) - p ≡ FADCB - 93

INTRODUCTION

L'analyse des modèles de régression amène les auteurs, dans une procédure commune, à examiner les erreurs résiduelles, afin de voir dans quelle mesure le modèle mathématique peut être mal formulé. Dans les études en séries chronologiques, les résidus autocorrelés peuvent être le résultat de la détermination incorrecte de la forme fonctionnelle de la partie fixe du modèle et/ou du fait que les erreurs suivent un processus autorégressif. Dans les études où les modèles utilisent les coupes instantanées, les résidus hétéroscédastiques sont dans de nombreux cas, le résultat de la première cause précitée et/ou du fait que les erreurs sont hétéroscédastiques. Dans le but d'éviter une erreur de spécification fonctionnelle, l'utilisation de la transformation paramétrique Box-Cox est suggérée. Des cas spéciaux de la famille paramétrique sont habituellement utilisés: formes linéaire, log-linéaire. Pour la partie stochastique, on postule des résidus en différences premières dans les modèles en séries chronologiques et une formulation arbitraire de l'hétéroscédasticité dans les modèles en coupes transversales. Par ailleurs on postulait jusqu'à tout récemment, la présence d'erreurs indépendantes et/ou homoscédastiques, si on estimait une forme fonctionnelle souple. Par la suite, la procédure Box-Cox a été généralisée aux erreurs qui suivent un processus autorégressif du premier ordre ou d'un ordre multiple.

Les diverses formes de l'hétéroscédasticité formulées dans la littérature sont des cas spéciaux d'une forme générale proposée par les Professeurs Gaudry et Dagenais (1978). L'application de la transformation paramétrique Box-Cox sur le terme d'erreur va nous permettre d'estimer la vraie forme de l'hétéroscédasticité en supposant qu'une ou deux variables d'activité est (sont) la source de l'hétéroscédasticité.

L'estimation conjointe de la forme fonctionnelle des parties fixe et stochastique, par la méthode du maximum de vraisemblance, va nous permettre de faire un discernement entre les deux forces qui agissent sur la forme fonctionnelle: la non-linéarité et l'hétéroscédasticité. Elle va également nous permettre de voir que les paramètres qui gouvernent la partie fixe, lesquels ne sont pas indépendants de ceux de la partie stochastique, affectent les élasticités de la demande, les signes des coefficients et les tests statistiques.

La minimisation de l'erreur ou encore la maximisation de la vraisemblance de l'échantillon nécessite d'établir à priori le moins de restrictions contraignantes et par conséquent la procédure Box-Cox est généralisée aux erreurs autocorrelées qui suivent un processus autorégressif d'un ordre multiple.

CHAPITRE I

ETUDE DE LA FORME FONCTIONNELLE DE LA PARTIE FIXE ET DE LA SPECIFICATION DE L'ERREUR: REVUE DE LA LITTERATURE.

L'étude systématique de la forme fonctionnelle flexible est un sujet relativement récent dans la littérature. L'application de transformations paramétriques Box-Cox sur les variables explicatives est alors utilisée et elle sert à minimiser l'erreur de spécification dûe à un choix arbitraire et par conséquent peut-être irrationnel de la forme fonctionnelle.

Ainsi, aucune propriété de comportement[1] (laquelle n'est pas sujette à une justification) n'est imposée, comme c'est le cas pour une forme fonctionnelle choisie arbitrairement. Aucune recherche pour établir une forme fonctionnelle précise (linéaire, log-linéaire, etc.) n'est ainsi faite et les résultats sont obtenus à partir d'une forme fonctionnelle qui résulte du processus d'estimation. Par ailleurs, dans la plupart des études courantes, il y a une nécessité d'établir différentes fonctions, ce que rend possible la paramétrisation, et de choisir à l'aide du processus d'estimation, la forme fonctionnelle qui donne le "meilleur" ajustement. La technique utilisée est alors celle du "scanning" ou "grille orientée". Dans d'autres études, une méthode de maximisation non-linéaire est utilisée. L'utilisation des valeurs des fonctions de vraisemblance qui résultent de l'estimation des modèles permet de trouver non pas le "vrai" modèle puisque la transformation paramétrique ne le permet pas, mais de trouver le "meilleur" modèle.

Plus récemment encore, la transformation paramétrique Box-Cox a été appliquée à la partie stochastique du modèle dans le but de déterminer une

---

1.- Une forme doublement logarithmique contraint à avoir des élasticités constantes.

forme fonctionnelle de l'hétéroscédasticité des erreurs beaucoup plus générale.

L'évolution de l'utilisation de la transformation Box-Cox sur les parties fixe et stochastique du modèle sera étudiée dans les deux dernières sections du premier chapitre. Cependant, avant de passer à ces points, nous allons présenter la spécification du terme résiduel telle qu'elle existe dans la théorie traditionnelle classique.

1.1. Cas classique.

Posons le problème sous une forme usuelle[2] où pour chaque observation nous avons:

$$y_t = \sum_k \beta_k X_{k_t} + u_t \quad , \quad \text{avec } t=1,\ldots,n, \qquad (1)$$

où n=nombre d'observations.

Ecrite sous une forme matricielle, l'expression (1) devient:

$$y = X\beta + u \quad , \qquad (2)$$

où y : est un vecteur d'observations sur la variable dépendante de dimension $(n \times 1)$,

X : est une matrice d'observations sur les k variables indépendantes fixées ou exogènes de dimension $(n \times k)$,

β : est un vecteur de paramètres de dimension $(k \times 1)$,

u : est un vecteur d'erreurs aléatoires de dimension $(n \times 1)$.

Nous émettons d'une manière implicite les hypothèses suivantes:

1. moyenne des erreurs est nulle, $E(u)=0$
2. matrice de variance-covariance[3], $E(uu')=\Omega=\psi^2 V$

2.- Nous ignorons le coefficient β_0 tout au cours de notre analyse afin de faciliter l'exposé.

3.- Lorsque V=I, c'est-à-dire lorsque $E(uu')=\psi^2 I$, nous postulons les hypothèses suivantes:

où V est une matrice symétrique définie positive d'ordre (n × n),

et

où ψ^2 est inconnu.

Il nous faut distinguer entre les éléments diagonaux et non-diagonaux de la matrice V.

1.1.1. Analyse de l'hétéroscédasticité du terme d'erreur.

L'analyse de l'hétéroscédasticité suggère que V est simplement une matrice diagonale dans laquelle les éléments diagonaux ne sont pas tous identiques (variance de l'erreur non constante).

Ainsi, la matrice Ω sera spécifiée comme étant égale à:

$$E(u_t u_{t'}) = E(u_t^2) = [f(X_{1_t},\ldots,X_{M_t})]\psi^2 \quad , \quad \text{pour tout } t=t' \qquad (3)$$

où $f(X_{1_t},\ldots,X_{M_t})$ est une fonction de plusieurs variables indépendantes,

$$= 0 \quad , \quad \text{pour tout } t\neq t'. \qquad (3')$$

Concernant la formulation mathématique de cette fonction, nous retrouvons dans la littérature diverses spécifications. Johnston (1972, pp.

---

†. $E(u_t^2) = \psi^2$, pour tout t, (hypothèse d'homoscédasticité de la variable aléatoire)

††. $E(u_t u_{t+s}) = 0$, pour $s\neq 0$, (hypothèse d'absence d'autocorrélation des erreurs résiduelles).

(I est une matrice d'identité d'ordre (n × n)).

Si ces deux dernières conditions ne sont pas respectées lorsque nous estimons par la méthode des moindres carrés ordinaires, ceci a pour conséquence de produire des écarts-types biaisés et des paramètres inefficaces bien que non-biaisés. Le postulat d'homoscédasticité est nécessaire (quoique non suffisant) si nous voulons obtenir des estimateurs à variance minimale, lorsque nous utilisons la méthode des moindres carrés ordinaires.

Si au moins un des deux postulats n'est pas respecté, les hypothèses sur la forme de la matrice variance-covariance Ω sont redéfinies (voir ci-dessus).

215-217) cite deux variantes de la forme de l'hétéroscédasticité pour lesquelles les éléments constituant la diagonale de V s'identifient à une variable explicative unique.

La formulation mathématique qui est la plus fréquemment utilisée est celle dans laquelle la variance de l'erreur est proportionnelle au carré de la variable X_k et elle est représentée par la relation explicite suivante:

$$E(u_t^2) = \psi^2 X_{k_t}^2. \tag{4}$$

La seconde formulation est représentée par l'égalité suivante:

$$E(u_t^2) = \psi^2 X_{k_t}. \tag{5}$$

La présence[4] de la variable X_{k_t} est identifiée par le test paramétrique de Goldfeld et Quandt (1965, pp.539-547).

En plus de ces formulations, diverses autres ont été élaborées par des auteurs tels que Glejser (1969,pp.316-323), Theil (1951, cité dans Goldfeld et Quandt, 1972, p.82).

Glejser propose des formulations qui sont plus ou moins générales mais qui se limitent à estimer l'hétéroscédasticité (test constructif) dans le cas d'une seule variable. Nous pouvons noter les diverses formes suivantes:

$$\text{(i)} \quad |u_t| = b_0 + b_1 X_{k_t}$$

$$\text{(ii)} \quad |u_t| = b_0 + b_1 X_{k_t}^{\frac{1}{2}} \tag{6}$$

4.- Outre ce test paramétrique de Goldfeld-Quandt (F test) nous pouvons noter d'autres tests de détection de la variable susceptible de causer l'hétéroscédasticité (à partir des résidus des moindres carrés ordinaires). En effet, parmi les tests qualifiés de tests non-constructifs (qui établissent la présence ou l'absence de l'hétéroscédasticité et ignorent l'aspect d'estimation) par Goldfeld et Quandt, nous retraçons dans la littérature, les études de Pearson et Barlett (test standard dans un contexte de non-régression en utilisant la méthode du rapport de vraisemblance), le test non-paramétrique (peak-test) de Goldfeld et Quandt, l'approche heuristique (dont l'approche du graphique des résidus fait partie), les travaux de Theil, etc. Ces études sont citées dans Goldfeld-Quandt (1972, pp.85-91).

(iii) $|u_t| = b_0 + b_1 X_{k_t}^{-1}$

etc.

La faiblesse apparente de cette approche provient de ce que nous ne connaissons pas à priori la spécification exacte des résidus. Aussi cette procédure par tâtonnements suggère d'estimer les diverses formes proposées et de retenir celle qui semble représenter le mieux l'hétéroscédasticité de la variable stochastique.

Pour Theil, la forme de l'hétéroscédasticité est telle que la variance de l'erreur est proportionnelle au carré de la moyenne de la variable dépendante. Elle est posée comme étant égale à:

$$E(u_t^2) = \gamma[E(y_t)]^2. \qquad (7)$$

1.1.2. Analyse de l'autocorrélation du terme d'erreur.

L'analyse de la corrélation entre les erreurs résiduelles suggère que les éléments hors-diagonaux de la matrice V sont différents de zéro.

Par conséquent, la matrice Ω sera spécifiée comme étant égale à:

$$E(u_t u_{t'}) = E(u_t^2) = \psi^2 \quad , \quad \text{pour } t=t' \qquad (8)$$

$$\neq 0 \quad , \quad \text{pour } t\neq t'. \qquad (8')$$

Nous supposons que u_t obéit à la structure d'autocorrélation suivante:

$$u_t = \sum_{\ell=1}^{r} \rho_\ell \, u_{t-\ell} + e_t \quad , \quad \text{pour } t=1,\ldots,n; \qquad (9)$$

- où les e_t sont des variables aléatoires distribuées identiquement et indépendamment avec

  1. moyenne des erreurs, $E(e_t) = 0$,

  2. matrice variance-covariance, $E(e_t e_{t'}) = \sigma^{**2} I$;

- où le processus satisfait à priori à la restriction de stationnarité standard, $|\Sigma\rho| < 1$.

Jusqu'ici, nous avons discuté de diverses formes d'hétéroscédasticité et d'une forme spécifique de la structure autorégressive. Celles-ci sont spécifiées et peuvent être modifiées indépendamment de la forme fonctionnelle de la partie fixe du modèle.

1.2. Analyse de la forme fonctionnelle de la partie fixe du modèle (avec $E(uu') = \psi^2 I$).

Lorsqu'il s'agit de déterminer la forme fonctionnelle de la partie fixe d'un modèle, il n'y a pas à priori de rationalité économique.

> "La théorie économique donne peu d'indication concernant les formes fonctionnelles appropriées à la spécification des relations économiques, et la présence des termes d'erreurs aléatoires dans les équations stochastiques ajoute un élément additionnel d'ambiguïté fonctionnelle." (Citation traduite de Dhrymes et autres, 1972, p.294).

En pratique, l'évidence empirique nous suggère habituellement de choisir (arbitrairement ou suite à des études antérieures effectuées dans le champ théorique du sujet à l'étude) entre une forme linéaire et une forme log-linéaire. Cependant comme le citent Gaudry et Dagenais (1978, p.1), nous pouvons nous poser la question de savoir si la transformation logarithmique est utilisée parce que la forme appropriée est vraiment log-linéaire ou si elle est utilisée parce qu'elle permet d'anticiper des résidus qui respectent approximativement l'hypothèse d'homoscédasticité.

L'alternative qui s'offre est donc d'estimer le modèle dans une forme aussi générale que possible afin de laisser les données elles-mêmes déterminer la forme de la fonction qui leur convient.

L'introduction systématique des considérations d'une forme fonctionnelle flexible dans les modèles de régression avec variables positives est rendue possible par l'utilisation de transformations monotoniques. Celles-ci ont été étudiées d'abord par Anscombe et Tukey (1954, cité dans Gaudry et Dagenais, 1978, p.3), Tukey (1957), Turner et autres (1961), Box et Tidwell (1962), Box et Cox (1964).

. Transformation Box-Cox.

Si nous considérons la transformation de puissance sur la variable dépendante et sur les variables indépendantes, l'équation (1) devient:

$$y_t^{(\lambda_y)} = \sum_k \beta_k X_{k_t}^{(\lambda_{xk})} + u_t \qquad (10)$$

où λ_y : est un scalaire,

λ_{xk} : est un élément du vecteur λ_x,

β_k : est un élément du vecteur β.

Si la variable dépendante y_t ou la k ième variable indépendante "non Bolean" X_{k_t} sont dénotées par Z_t, la transformation Box-Cox [5] de chaque variable Z_t est définie par le paramètre λ comme étant

$$Z_t^{(\lambda)} = \begin{cases} \dfrac{Z_t^{\lambda} - 1}{\lambda} & , \quad \lambda \neq 0 \\ \text{ℓn } Z_t \text{ [6]} & , \quad \lambda \doteq 0. \end{cases} \qquad (11)$$

5.- La transformation Box-Cox est un cas spécial de la transformation Box-Tukey, laquelle est définie comme étant

$$(Z_t + \mu)^{(\lambda)} = \begin{cases} \dfrac{(Z_t + \mu)^{\lambda} - 1}{\lambda} & , \quad \lambda \neq 0, \quad Z_t + \mu > 0 \\ \text{ℓn } (Z_t + \mu) & , \quad \lambda \doteq 0, \quad Z_t + \mu > 0. \end{cases}$$

Dans la transformation Box-Cox, nous contraignons le paramètre de localisation μ à être égal à zéro.

6.- Par la règle de l'Hôpital, il est possible de démontrer que

Lorsque λ est égal à 1, nous parlons en termes de forme fonctionnelle linéaire et lorsque λ tend vers 0, nous faisons référence à la forme fonctionnelle logarithmique. Si λ est égal à -1, il en résulte la forme fonctionnelle réciproque, etc.

Schlesselman (1971, p.308) a souligné l'importance d'inclure une constante additive dans le modèle de régression dans le but de préserver l'invariance des transformations aux unités de mesure.

La transformation de puissance inverse de la définition précitée est[7]:

$$Z_t = \begin{cases} (\lambda Z_t^{(\lambda)} + 1)^{1/\lambda} & , \quad \lambda \neq 0 \\ \exp\,(Z_t^{(\lambda)}) & , \quad \lambda \to 0. \end{cases} \qquad (11')$$

Certains auteurs tels Heckman et Polachek (1974), Kau et Sirmans (1976), Welland[8] (1976), Spitzer (1976, 1977: modèle à équations simultanées) ont modifié les variables originales utilisées dans leur(s) modèle(s) de régression en introduisant la transformation de puissance directe. Cependant, ils n'ont pas examiné l'impact de cette transformation sur la partie stochastique du modèle. Zarembka (1968, 1974) a, en sus, postulé une matrice variance-covariance différente de $\psi^2 I$, $[E(u_t u_{t'}) \neq 0$, pour $t \neq t']$ en considérant l'autocorrélation du premier ordre, mais il n'a pas estimé son modèle en tenant compte simultanément de la structure de la partie stochastique et de la

---

$\lim_{\lambda \to 0} \frac{Z_t^{\lambda} - 1}{\lambda} \to \ln Z_t$ de sorte que la fonction est continue à $\lambda=0$ (Box-Cox, 1964, pp.214-215).

7.- Inverse du log = exponentielle.

8.- Les sources des études de Kau et Sirmans, Welland sont citées dans Gaudry et Dagenais (1978, p.3).

partie fixe du modèle.

Nous allons, à présent, étendre le problème en ajoutant des raffinements additionnels à la partie aléatoire du modèle et étudier la spécification conjointe des parties fixe et stochastique du modèle.

Les éléments théoriques qui seront discutés dans la prochaine sous-section, représentent le point de départ majeur de notre discussion.

1.3. Analyse de la forme fonctionnelle de la partie fixe simultanément avec la partie stochastique du modèle (avec $E(uu') \neq \psi^2 I$).

La transformation des variables originales par des paramètres de puissance affecte la distribution des erreurs. Si en théorie un des objectifs de cette transformation est de normaliser la distribution de l'erreur [$u \sim N(0, \psi^2)$], en pratique cet objectif n'est pas toujours atteint. Une autre difficulté avec la transformation Box-Cox est que le terme d'erreur est supposé avoir une variance constante à travers les observations[9]. Cependant, ceci peut ne pas être le cas. Nonobstant l'importance de tenir compte des problèmes que posent la non-normalité et la variance non-constante du terme d'erreur, nous constatons que l'intérêt porté aux effets de la forme fonctionnelle sur la distribution des résidus est relativement récent dans la littérature statis-

---

9.- Tukey (1957, p.609) a noté que le but de la transformation est d'augmenter le degré d'approximation, à laquelle les trois propriétés suivantes sont associées:

†. les effets sont additifs (la transformation des variables peut amener un modèle à tendre vers le linéaire: Zarembka, 1974, p.84),
††. la variabilité de l'erreur est constante (stabiliser la variance de l'erreur), et/ou
†††. la distribution de l'erreur est symétrique et possiblement presque normale.

D'après Zarembka (pp.87-88), les transformations des variables indépendantes sont concentrées seulement sur la première propriété puisque ces variables sont concentrées sur la valeur de la variable dépendante. D'un autre côté, il mentionne que la variable dépendante est aléatoire et que les deux hypothèses sur le terme d'erreur sont reliées à cette variable.

tique.

Nous devons à Zarembka (1974) une première approche sur la détermination conjointe de l'hétéroscédasticité et de la forme fonctionnelle de la partie fixe. Ce dernier a proposé une mesure[10] approximative de la sensibilité de la forme fonctionnelle à la présence éventuelle d'une forme particulière d'hétéroscédasticité dans le terme d'erreur.

Par la suite, Major (1978) a énoncé une formulation de l'hétéroscédasticité[11] et l'a estimée conjointement avec la forme de la partie fixe du modèle (pour différentes valeurs de lambdas donnés).

Finalement, Gaudry et Dagenais (1978) ont formulé une spécification plus générale de la forme de l'hétéroscédasticité et ont montré que celle-ci peut être analysée simultanément avec la forme de la partie fixe du modèle.

Les paramètres λ peuvent être plus sensibles à la présence de l'hétéroscédasticité qu'à la présence de l'autocorrélation des résidus. Si dans le cas linéaire, nous supposons de façon générale, que le postulat d'homoscédasticité est satisfait dans les régressions en séries chronologiques (étant donné que les variations de la variable dépendante sont moins importantes que dans les régressions en coupes instantanées), dans le cas non-linéaire, nos anticipations sont modifiées.[12] Les transformations Box-Cox peuvent introduire l'hé-

---

10.- La mesure d'approximation développée par cet auteur est présentée dans Zarembka, p.94.

11.- La formulation proposée par Major s'exprime de la manière suivante:

$$E(u_t^2) = \sigma^2 \exp(\delta_1 X_{1_t} + \delta_2 X_{2_t} + \ldots + \delta_k X_{k_t})\, v$$

où $E(v) = 0$ et $E(v^2) =$ constante.

12.- Si nous posons que $\lambda=2$, par exemple, le rapport des variances des grandes valeurs sur les petites valeurs devient plus grand et le problème de l'hétéroscédasticité peut survenir.

téroscédasticité ou peuvent également, tel que mentionné précédemment, changer le caractère hétéroscédastique des résidus, c'est-à-dire qu'elles peuvent compenser pour une mauvaise forme de la partie aléatoire du modèle. (L'estimation par le maximum de vraisemblance de la forme fonctionnelle peut régulariser la distribution des résidus (Gaudry et Wills, 1978, p.32)).

Dans leur étude, Gaudry et Dagenais (1978) ont étendu leur analyse en généralisant la procédure Box-Cox aux erreurs qui suivent un processus d'autocorrélation multiple.

Nous pouvons noter que Savin et White (1978) furent les premiers auteurs à analyser la forme fonctionnelle de la partie fixe d'un modèle conjointement avec la structure de sa partie stochastique en considérant un schéma autorégressif du premier ordre (estimation simultanée de λ et ρ_1). Gaudry et Wills (1978) ont étendu le problème à la structure d'autocorrélation multiple.

Nous allons présenter maintenant la forme générale de l'hétéroscédasticité telle que spécifiée par Gaudry et Dagenais.

1.3.1. Forme de l'hétéroscédasticité dans un modèle linéaire.

1.3.1.1. Une forme générale pour les erreurs hétéroscédastiques.

Supposons un modèle linéaire dans lequel la matrice variance-covariance des u est posée comme étant

$$E(uu') = \Omega = \psi^2 \begin{bmatrix} f(X_1) & 0 & \cdots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & f(X_t) & 0 \\ \vdots & & & \ddots \\ 0 & \cdots & 0 & f(X_n) \end{bmatrix} \qquad (12)$$

avec

$$f(X_t) = \{\lambda_u [\delta_o + \sum_m \delta_m X_{m_t}^{(\lambda_{em})}] + 1\}^{1/\lambda_u} > 0 \quad \underline{\text{si } \lambda_u \neq 0} \qquad (13)$$

où

$X_{m_t}^{(\lambda_{em})}$: représente la variable X_{m_t} élevée à la puissance λ_{em},

X_{m_t} : représente la variable explicative qui est susceptible de causer de l'hétéroscédasticité parmi les k variables indépendantes (13), à la période t,

λ_{em} et δ_m : sont respectivement des éléments typiques des vecteurs λ_e et δ,

λ_u et δ_o : sont des scalaires.

Nous constatons que les transformations Box-Cox directes sont appliquées aux X_m variables explicatives de la variance de l'erreur et qu'une transformation de puissance inverse est appliquée au contenu des parenthèses carrées. Les constantes δ_o et ψ sont introduites car autrement nous aurions des procédures qui ne seraient pas invariantes aux unités de mesure des X et de y.

Compte tenu des différentes valeurs qui sont attribuées à λ_u et à λ_{em}, l'équation (13) peut prendre plusieurs formes. Ainsi,

. <u>si $\lambda_u = 1$</u>, l'équation (13) devient après multiplication par ψ^2, égale à:

$$\tilde{\tilde{w}}_{tt} = \tilde{\tilde{\delta}}_o + \sum_m \tilde{\tilde{\delta}}_m X_{m_t}^{(\lambda_{em})}$$

$$\text{où } \tilde{\tilde{\delta}}_o = (\delta_o + 1)\psi^2 \qquad (14)$$

$$\tilde{\tilde{\delta}}_m = \delta_m \psi^2 .$$

13.- Nous pouvons noter qu'en sus des variables explicatives de la partie fixe $(X_1,\ldots,X_M)$, nous pouvons ajouter d'autres variables exogènes $(Z_1,\ldots. Z_{M)}$ lesquelles Z_m variables peuvent différer des X_k variables de la partie fixe

Selon la valeur que prend λ_{em}, nous avons une forme linéaire si λ_{em} = 1 pour tout m et plusieurs cas spéciaux de formes non-linéaires: racine carrée si λ_{em} = 1/2 et puissance carrée si λ_{em} = 2. Par ailleurs,

. si λ_u = 0, nous avons à partir de la condition (11'):

$$f(X_t) = \exp\left(\delta_0 + \sum_m \delta_m X_{m_t}^{(\lambda_{em})}\right) . \qquad (15)$$

Alors l'équation (15) devient après multiplication par ψ^2, égale à:

$$\begin{aligned} \tilde{w}_{tt} &= \psi^2 \exp\left(\delta_0 + \sum_m \delta_m X_{m_t}^{(\lambda_{em})}\right) \\ &= \psi^2 \exp(\delta_0) \cdot \exp\left(\sum_m \delta_m X_{m_t}^{(\lambda_{em})}\right) \\ &\equiv \sigma^2 \left[\exp\left(\sum_m \delta_m X_{m_t}^{(\lambda_{em})}\right)\right] . \end{aligned} \qquad (16)$$

En posant λ_{em} = 0 pour tout m, nous obtenons une forme multiplicative multivariée, du type Dagum et Dagum (1974)[14],

$$\hat{w}_{tt} = \sigma^2 \prod_m X_{m_t}^{\delta_m} . \qquad (17)$$

14.- Si λ_{em} = 0 pour tout m, l'équation (16) s'écrit, compte tenu de la condition (11), de la façon suivante:

$$\begin{aligned} \hat{w}_{tt} &= \sigma^2 \left[\exp\left(\sum_m \delta_m \ln X_{m_t}\right)\right] \\ &= \sigma^2 \left[\exp(\delta_1 \ln X_{1_t}) \exp(\delta_2 \ln X_{2_t}) \ldots \exp(\delta_m \ln X_{m_t})\right] \\ &= \sigma^2 \left[\exp(\ln X_{1_t}^{\delta_1}) \exp(\ln X_{2_t}^{\delta_2}) \ldots \exp(\ln X_{m_t}^{\delta_m})\right] \end{aligned}$$

et comme $\exp(\ln X) = X$, nous avons

$$\begin{aligned} \hat{w}_{tt} &= \sigma^2 \left[X_{1_t}^{\delta_1} , X_{2_t}^{\delta_2} , \ldots , X_{m_t}^{\delta_m}\right] \\ &= \sigma^2 \prod_m X_{m_t}^{\delta_m} . \end{aligned}$$

Cette forme multiplicative, qui utilise une fonction Cobb-Douglas, ne répond pas en totalité au problème de l'estimation de l'hétéroscédasticité.

Si $\delta_m = 2$ pour tout m, nous avons une forme multiplicative classique à plusieurs variables,

$$\hat{w}_{tt} = \sigma^2 \prod_m X_{m_t}^2 \quad . \tag{18}$$

En posant tous les δ_m de l'expression (16) égalent à zéro sauf un et en posant le λ_{em} de la variable restante égale à zéro, nous obtenons la spécification de Park (1966),

$$\ddot{w}_{tt} = \sigma^2 X_{m_t}^{\delta_m} \quad . \tag{19}$$

En imposant la contrainte que $\delta_m = 2$, nous obtenons la formulation classique telle que déjà spécifiée,

$$\dot{w}_{tt} = \sigma^2 X_{m_t}^2 \quad . \tag{20}$$

1.3.1.2. La fonction de vraisemblance pour un modèle linéaire.

Si nous considérons le modèle $y = X\beta + u$, où u est supposé être distribué normalement avec un vecteur de moyenne nulle et une matrice de covariance (12) dans laquelle les éléments diagonaux sont spécifiés comme dans (15)[15], la fonction de vraisemblance de l'échantillon peut s'écrire comme étant

---

15.- Gaudry et Dagenais proposent d'utiliser dans la dérivation de la fonction de vraisemblance, la spécification telle que définie dans (16) et ce pour deux raisons. Primo, la sous-classe définie dans (14) peut produire des éléments négatifs dans Ω. Secundo, la sous-classe définie dans (16) permet des interactions entre les variables (un petit sous-ensemble des k variables explicatives) les plus probables à expliquer la variance de l'erreur. Aussi, dans la poursuite de notre analyse, nous retiendrons la spécification (16).

égale à[16]:

$$\mathcal{L}(y;\ \beta,\ \sigma^2,\ \lambda_e,\ \delta) = \left(\frac{1}{\sqrt{2\pi}}\right)^n |\tilde{\Omega}|^{-1/2} \exp\left\{ -\frac{1}{2}(y - X\beta)' \tilde{\Omega}^{-1} (y - X\beta) \right\}. \quad (21)$$

Si la relation $y = X\beta + u$ est posée sous une forme usuelle et si Ω est écrit sous sa forme explicite (16), alors l'équation (21) devient:

$$\mathcal{L}(y;\ \beta_k,\ \sigma^2,\ \lambda_{em},\ \delta_m) = \prod_{t=1}^{n} \frac{1}{\sqrt{2\pi}\ \sigma \sqrt{\exp(\Sigma_m \delta_m X_{m_t}^{(\lambda_{em})})}} \exp\left\{ -\frac{(y_t - \Sigma_k \beta_k X_{k_t})^2}{2\sigma^2 \exp(\Sigma_m \delta_m X_{m_t}^{(\lambda_{em})})} \right\}. \quad (22)$$

Cette fonction est donc maximisée par rapport aux β_k de la partie fixe, à σ et aux paramètres inclus dans $f(X_t)$ à savoir, λ_{em} et δ_m.

1.3.1.3. La fonction de vraisemblance pour un modèle non-linéaire.

Considérons maintenant la fonction de vraisemblance pour un modèle non-linéaire.

Si nous introduisons des transformations paramétriques Box-Cox sur la partie du modèle linéaire et que $f(X_t)$ (pour $t=1,\ldots,n$) dans la matrice de covariance (12) est défini comme dans (15), alors la fonction de vraisemblance est ainsi formulée:

$$\mathcal{L}(y^{(\lambda_y)};\ \beta,\ \sigma^2,\ \lambda_x,\ \lambda_e,\ \delta) = \left(\frac{1}{\sqrt{2\pi}}\right)^n |\tilde{\Omega}|^{-1/2} \exp\left\{ -\frac{1}{2}(y^{(\lambda_y)} - X^{(\lambda_x)}\beta)' \tilde{\Omega}^{-1} (y^{(\lambda_y)} - X^{(\lambda_x)}\beta) \right\}. \quad (23)$$

Comme y est observé, la vraisemblance de l'échantillon devient:

16.- Le vecteur δ est constitué par les éléments δ_m puisque nous savons que δ_0 est inclus dans σ^2.

$$\mathcal{L}(y;\ \beta,\ \sigma^2,\ \lambda_y,\ \lambda_x,\ \lambda_e,\ \delta) = \left(\frac{1}{\sqrt{2\pi}}\right)^n |\tilde{\Omega}|^{-1/2} \cdot$$

$$\exp\left\{ -\frac{1}{2} (y^{(\lambda_y)} - X^{(\lambda_x)}\beta)' \tilde{\Omega}^{-1} (y^{(\lambda_y)} - x^{(\lambda_x)}\beta) \right\} |J(\lambda_y;y)| , \qquad (24)$$

où la valeur absolue du Jacobien[17] de la transformation de $y_t^{(\lambda_y)}$ aux valeurs observées de y_t:

$$|J(\lambda_y;y)| = \left| \det \frac{\partial y_t^{(\lambda_y)}}{\partial y_t} \right| = \prod_{t=1}^{n} y_t^{\lambda_y - 1} . \qquad (25)$$

L'expression (24) peut être réécrite comme:

$$\mathcal{L}(y;\ \beta_k,\ \sigma^2,\ \lambda_y,\ \lambda_{xk},\ \lambda_{em},\ \delta_m) =$$

$$\prod_{t=1}^{n} \frac{y_t^{\lambda_y - 1}}{\sqrt{2\pi}\ \sigma \sqrt{\exp(\Sigma_m \delta_m X_{m_t}^{(\lambda_{em})})}} \exp\left\{ -\frac{(y_t^{(\lambda_y)} - \Sigma_k \beta_k X_{k_t}^{(\lambda_{xk})})^2}{2\sigma^2 \exp(\Sigma_m \delta_m X_{m_t}^{(\lambda_{em})})} \right\} . \qquad (26)$$

En ajoutant aux paramètres inclus dans la forme linéaire (σ, λ_{em}, δ_m), les paramètres (λ_y, λ_{x1}, ... , λ_{xk}), nous obtenons une formulation qui nous permet de considérer simultanément la forme de la partie fixe du modèle et la forme de l'hétéroscédasticité.

Nous allons aborder maintenant le cas d'une forme fonctionnelle générale pour laquelle la matrice variance-covariance des erreurs résiduelles Ω est définie de telle sorte qu'elle tient compte conjointement d'une forme d'hétéroscédasticité (telle que spécifiée dans 15) et d'une structure d'autocorrélation, soit du premier ordre, soit d'un ordre multiple.

17.- Si $\lambda_y = 1$ (cas linéaire), $|J(\lambda_y;y)| = 1$.

1.3.2. Formulation complète d'un modèle non-linéaire en présence de l'hétéroscédasticité et d'une structure d'autocorrélation.

1.3.2.1. du premier ordre.

La prise en considération simultanée d'un processus autorégressif et d'un terme stochastique hétéroscédastique nous amène à reformuler les hypothèses du modèle initial. Dans le cas d'un modèle non-linéaire, la spécification devient:

$$y_t^{(\lambda_y)} = \sum_k \beta_k X_{k_t}^{(\lambda_{xk})} + u_t \quad , \tag{27}$$

avec

$$u_t = \sqrt{f(X_t)}\, v_t \quad . \tag{28}$$

Si le terme aléatoire v_t est généré par un schéma autorégressif du premier ordre, il nous faut introduire la relation suivante:

$$v_t = \rho_1 v_{t-1} + s_t \quad , \quad \text{où } |\rho_1| < 1 \; . \tag{29}$$

Nous supposons implicitement que v_t et s_t respectent successivement les hypothèses suivantes:

$$\cdot\ E(v_t v_{t'}) = E(v_t^2) = \psi_s^2/(1 - \rho_1^2) = \psi^2 \quad , \quad \text{pour } t=t'$$
$$\neq 0 \quad , \quad \text{pour } t\neq t' \; .$$

(Puisque nous postulons ici seulement l'autocorrélation du premier ordre, nous posons si t'=t-1, que $E(v_t v_{t-1}) = \rho_1 \psi^2$).

Et s_t est normalement et indépendamment distribué avec:

$$\cdot\ E(s_t) = 0$$
$$\cdot\ E(s_t s_{t'}) = E(s_t^2) = \psi^{*2} = \psi^2 (1 - \rho_1^2) \quad , \quad \text{pour } t=t'$$
$$= 0 \quad , \quad \text{pour } t\neq t' \; .$$

Si le processus d'erreur est stationnaire, le vecteur d'erreurs s

peut être écrit, tel que chez Goldfeld et Quandt (1972, p.184), Kadiyala (1968, p.94) et autres, dans la forme suivante:

$$s = Tv \; . \tag{30}$$

La matrice de transformation T d'ordre $(n \times n)$ est donnée par:

$$T = \begin{bmatrix} \sqrt{1-\rho_1^2} & 0 & \ldots & 0 & 0 \\ -\rho_1 & 1 & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & \ldots & \ldots & 1 & 0 \\ 0 & 0 & \ldots & -\rho_1 & 1 \end{bmatrix} \tag{31}$$

et s_1 correspond par conséquent à $v_1 \sqrt{1 - \rho_1^2}$.

Après avoir énoncé les hypothèses sur v_t et sur s_t, et noté la transformation appropriée qui va nous permettre de transformer l'équation (27) en un modèle de régression standard, nous allons maintenant énoncer celles sur u.

Posons $u = HT^{-1}s$

où $H = \text{diagonale} \; [(f(X_t))^{1/2}]$.

Alors la matrice variance-covariance des u est égale à:

$$\begin{aligned} E(uu') &= E(HT^{-1}ss'(T^{-1})'H') \\ &= HT^{-1}(T^{-1})'H'E(ss') \\ &= HT^{-1}(T^{-1})'H'\psi^{*2} = \Omega^* \; . \end{aligned}$$

1.3.2.1.1. La fonction de vraisemblance en présence de l'hétéroscédasticité et de l'autocorrélation du premier ordre.

Si nous introduisons simultanément les spécifications d'hétéroscédasticité et d'autocorrélation telles que formulées dans (28) et (29) et si nous postulons que $\lambda_u = 0$ et que $f(X_t)$ est défini comme dans (15), la fonction de

vraisemblance (24) s'écrit de la manière suivante:

$$\mathcal{L}(y;\ \beta,\ \sigma^{*2},\ \rho_1,\ \lambda_y,\ \lambda_x,\ \lambda_e,\ \delta) = \left(\frac{1}{\sqrt{2\pi}\ \sigma^*}\right)^n |\tilde{H}|^{-1/2} |T| \cdot$$

$$\exp\left\{ -\frac{1}{2} (y^{(\lambda_y)} - X^{(\lambda_x)}\beta)' \tilde{\Omega}^{*-1} (y^{(\lambda_y)} - X^{(\lambda_x)}\beta) \right\} |J(\lambda_y;y)| , \qquad (32)$$

où[18] $\tilde{\Omega}^* = \tilde{H}T^{-1}(T^{-1})'\tilde{H}'\sigma^{*2}$,

avec $\sigma^{*2} = \sigma^2 (1 - \rho_1^2)$

$\tilde{H}$ = diagonale $[\exp(\Sigma_m \delta_m X_{m_t}^{(\lambda_{em})}/2)]$.

Puisque le Jacobien de la transformation de s à v, correspond au déterminant $T = (1 - \rho_1^2)^{1/2}$, la fonction de vraisemblance (32) correspond à:

$$\mathcal{L}(y;\ \beta_k,\ \sigma^{*2},\ \rho_1,\ \lambda_y,\ \lambda_{xk},\ \lambda_{em},\ \delta_m) = (1 - \rho_1^2)^{1/2} \frac{y_1^{\lambda_y - 1}}{\sqrt{2\pi}\ \sigma^* \sqrt{\exp(\Sigma_m \delta_m X_{m_1}^{(\lambda_{em})})}} .$$

18.- Nous avons posé ci-dessus que H = diagonale $[(f(X_t))^{1/2}]$.

Si $\lambda_u = 0$, nous avons H = diagonale $[\exp(\delta_o + \Sigma_m \delta_m X_{m_t}^{(\lambda_{em})}/2)]$, où le terme entre [ ] vient de ce que $[e^{\alpha}]^{1/2} = [e^{\alpha/2}]$.

Alors $\psi^{*2} H = \psi^{*2}$ diagonale $[\exp(\delta_o + \Sigma_m \delta_m X_{m_t}^{(\lambda_{em})}/2)]$

$= \psi^2 (1 - \rho_1^2)\ e^{\delta_o}$ diagonale $[\exp(\Sigma_m \delta_m X_{m_t}^{(\lambda_{em})}/2)]$ et puisque

$\psi^2 e^{\delta_o} = \sigma^2$, nous avons:

$\psi^{*2} H = \sigma^2 (1 - \rho_1^2)$ diagonale $[\exp(\Sigma_m \delta_m X_{m_t}^{(\lambda_{em})}/2)]$

$= \sigma^{*2} \tilde{H}$.

$$\exp\left\{ - \frac{(1-\rho_1^2)\,(y_1^{(\lambda_y)} - \sum_k \beta_k X_{k_1}^{(\lambda_{xk})})^2}{2\sigma^{*2}\exp(\sum_m \delta_m X_{m_1}^{(\lambda_{em})})} \right\} \cdot \prod_{t=2}^{n} \frac{y_t^{\lambda_y - 1}}{\sqrt{2\pi}\,\sigma^* \sqrt{\exp(\sum_m \delta_m X_{m_t}^{(\lambda_{em})})}} \cdot$$

$$\exp\left\{ - \frac{1}{2\sigma^{*2}} \left[\frac{(y_t^{(\lambda_y)} - \sum_k \beta_k X_{k_t}^{(\lambda_{xk})})}{\sqrt{\exp(\sum_m \delta_m X_{m_t}^{(\lambda_{em})})}} - \rho_1 \frac{(y_{t-1}^{(\lambda_y)} - \sum_k \beta_k X_{k_{t-1}}^{(\lambda_{xk})})}{\sqrt{\exp(\sum_m \delta_m X_{m_{t-1}}^{(\lambda_{em})})}} \right]^2 \right\} \cdot (33)$$

Les termes qui se réfèrent à la première observation, sont isolés de ceux qui se réfèrent aux autres observations.

1.3.2.1.2. Une formulation simplifiée de la vraisemblance en présence de l'autocorrélation du premier ordre.

Afin d'éviter des complications[19] introduites par la première observation, nous pouvons éliminer celle-ci et maximiser la vraisemblance sur les n-1 dernières observations.

Beach et MacKinnon (1978, p.52), tout comme Goldfeld et Quandt (1972, p.185), affirment qu'asymptotiquement, le fait d'ignorer la première observation n'implique aucune différence. Les premiers ont cependant souligné que dans les petits échantillons, cette omission peut entraîner une différence substantielle[20].

19.- Si nous incluons la première observation et que $|\rho_1|$ tende vers un, alors le premier terme de (33) va tendre vers zéro et par conséquent, il en sera de même pour la fonction de vraisemblance.

20.- En considérant un modèle linéaire homoscédastique, avec une constante et une variable explicative pour des échantillons de 20 et 50 observations, Beach et MacKinnon (1978, pp.54-56) ont présenté des résultats d'une analyse de Monte Carlo pour différentes valeurs de ρ. Les résultats obtenus suggèrent que dans plusieurs cas, des différences significatives peuvent être relevées (en termes de biais, de la racine de la moyenne des erreurs au carré) selon qu'ils utilisent la technique conventionnelle de Cochrane-Orcutt, ou qu'ils utilisent une technique alternative du maximum de vraisemblance (FML) qui incorpore la première observation. Les gains d'inclure la première observation

D'après eux, il serait préférable, d'un point de vue purement théorique d'utiliser une procédure du maximum de vraisemblance qui ignore l'observation initiale mais qui incorpore à priori dans la fonction de vraisemblance elle-même, la condition de stationnarité $|\rho| < 1$.

Si nous excluons la première observation dans le cas d'un schéma autorégressif du premier ordre, ceci implique que nous ignorons a) le premier terme du Jacobien, à savoir $y_1^{\lambda_y - 1}$ dans l'expression (33) b) u_1^2 c) la première rangée de T dans (31). Nous avons donc

$$T_1 = \begin{bmatrix} -\rho_1 & 1 & 0 \ldots 0 & 0 \\ 0 & -\rho_1 & 1 \ldots 0 & 0 \\ \ldots\ldots\ldots\ldots\ldots & & & \\ 0 & 0 & 0 \ldots 1 & 0 \\ 0 & 0 & 0 \ldots -\rho_1 & 1 \end{bmatrix}$$

et la valeur du déterminant de T est maintenant 1 au lieu de $(1 - \rho_1^2)^{1/2}$.

La fonction de vraisemblance (33) se transforme en:

$$\mathcal{L}(y;\ \beta_k,\ \sigma^{*2},\ \rho_1,\ \lambda_y,\ \lambda_{xk},\ \lambda_{em},\ \delta_m) =$$

$$\prod_{t=2}^{n} \frac{[y_t^{\lambda_y - 1}]\sim}{\sqrt{2\pi}\,\sigma^*} \exp\left\{ - \frac{1}{2\sigma^{*2}} \left([y_t^{(\lambda_y)}]\tilde{\star} - \sum_k \beta_k [X_{k_t}^{(\lambda_{xk})}]\tilde{\star}\right)^2 \right\} \qquad (34)$$

avec

$$[\xi_t]\sim = \frac{\xi_t}{\sqrt{\exp(\sum_m \delta_m X_{m_t}^{(\lambda_{em})})}} \quad \text{et} \quad [\xi_t]\tilde{\star} = \frac{\xi_t}{\sqrt{\exp(\sum_m \delta_m X_{m_t}^{(\lambda_{em})})}} - \rho_1 \frac{\xi_{t-1}}{\sqrt{\exp(\sum_m \delta_m X_{m_{t-1}}^{(\lambda_{em})})}} . \qquad (35)$$

---

sont supérieurs, contrairement à leurs anticipations, dans l'échantillon de 50 observations; aussi d'après eux, il faut un très grand échantillon pour rendre les deux procédures également efficaces.

1.3.2.1.3. Une formulation simplifiée de la vraisemblance en présence d'une structure d'autocorrélation multiple.

En présence d'une structure d'autocorrélation multiple[21] ainsi définie

$$v_t = \sum_{\ell=1}^{r} \rho_\ell \, v_{t-\ell} + e_t \quad , \tag{29'}$$

la fonction de vraisemblance sera simplifiée en éliminant autant d'observations que l'ordre d'autocorrélation le plus élevé considéré le nécessite. Cette exclusion permet d'éviter les problèmes considérables d'estimation que susciterait une formulation complète de ladite fonction de vraisemblance.

La fonction de vraisemblance exprimée dans (34) se ramène à:

$$\mathcal{L}(y;\ \beta_k,\ \sigma^{**2},\ \rho_\ell,\ \lambda_y,\ \lambda_{xk},\ \lambda_{em},\ \delta_m) =$$

$$\prod_{t=1+r}^{n} \frac{[y_t^{\lambda_y - 1}]\sim}{\sqrt{2\pi}\ \sigma^{**}} \exp\left\{ - \frac{1}{2\sigma^{**2}} \left([y_t^{(\lambda_y)}]\tilde{}\tilde{} - \sum_k \beta_k [X_{k_t}^{(\lambda_{xk})}]\tilde{}\tilde{}\right)^2 \right\} \tag{36}$$

où r = ordre d'autocorrélation le plus élevé considéré.

Sous les conditions habituelles de stationnarité, le vecteur ρ sera égal à $\rho_1, \ldots, \rho_\ell, \ldots, \rho_r$ de telle sorte que $[\xi_t]\tilde{}\tilde{}$ défini dans (35) sera remplacé par:

$$[\xi_t]\tilde{}\tilde{} = \frac{\xi_t}{\sqrt{\exp(\sum_m \delta_m X_{m_t}^{(\lambda_{em})})}} - \sum_{\ell=1}^{r} \rho_\ell \frac{\xi_{t-\ell}}{\sqrt{\exp(\sum_m \delta_m X_{m_{t-\ell}}^{(\lambda_{em})})}} \quad , \tag{37}$$

$[\xi_t]\sim$ est défini comme dans (35) et σ^{**} est plus grand que zéro.

21.- En présence d'une structure d'autocorrélation multiple, nous remplaçons l'équation (29) par la relation (29').

Le logarithme népérien de la fonction de vraisemblance est:

$$L(y;\ \beta_k,\ \sigma^{**2},\ \rho_\ell\ ,\ \lambda_y,\ \lambda_{xk},\ \lambda_{em},\ \delta_m) =$$

$$L(\omega) = (\lambda_y - 1) \sum_{t=1+r}^{n} \ln y_t - \left(\frac{n-r}{2}\right) \ln \sigma^{**2} - \frac{1}{2} \sum_{t=1+r}^{n} \ln f(X_t)$$

$$- \frac{1}{2\sigma^{**2}} \sum_{t=1+r}^{n} [y_t^{**} - \Sigma\beta_k X_{k_t}^{**}]^2 \qquad (38)$$

où $y_t^{**} - \Sigma_k \beta_k X_{k_t}^{**} = e_t$ (39)

et $y_t^{**} = y_t^{*} - \Sigma_\ell \rho_\ell\ y_{t-\ell}^{*}$ (40)

$$X_{k_t}^{**} = X_{k_t}^{*} - \Sigma_\ell \rho_\ell\ X_{k_{t-\ell}}^{*} \qquad (41)$$

$$y_t^{*} = \frac{y_t^{(\lambda_y)}}{\sqrt{f(X_t)}} \qquad (42)$$

$$X_{k_t}^{*} = \frac{X_{k_t}^{(\lambda_{xk})}}{\sqrt{f(X_t)}} \qquad (43)$$

et $f(X_t)$ est défini comme dans (15).

CHAPITRE II

PROCEDURE D'ESTIMATION.

L'aspect théorique tel que présenté dans le premier chapitre de cette étude nous confronte à un problème de maximisation[1] de fonctions non-linéaires. La position à adopter dans une telle situation est de maximiser la fonction de vraisemblance appropriée ou encore de minimiser la somme des résidus au carré.

La technique de maximisation suggère de maximiser le logarithme de la fonction de vraisemblance concentrée. Aussi dans le but de diminuer le nombre de paramètres impliqués dans le processus de maximisation, le logarithme de la fonction de vraisemblance (38) sera concentré sur σ^{**2} et sur les β_k.

Le logarithme de la fonction de vraisemblance concentrée sur σ^{**2} et sur les β_k est[2]:

$$L(y;\ \rho,\ \lambda_y,\ \lambda_x,\ \lambda_e,\ \delta) =$$

$$L(\omega^{**}) = (\lambda_y - 1) \sum_{t=1+r}^{n} \ell n\ y_t - \left\langle \frac{n-r}{2} \right\rangle \ell n\ \hat{\sigma}^{**2}(\rho,\lambda_y,\lambda_x,\lambda_e,\delta)$$

$$- \frac{1}{2} \sum_{t=1+r}^{n} \ell n\ f(X_t) \quad , \tag{44}$$

avec $\hat{\sigma}^{**2}(\rho,\ \lambda_y,\ \lambda_x,\ \lambda_e,\ \delta) =$

$$\frac{1}{(n-r)} [y^{**} - X^{**}\hat{\beta}(\rho,\lambda_y,\lambda_x,\lambda_e,\delta)]' [y^{**} - X^{**}\hat{\beta}(\rho,\lambda_y,\lambda_x,\lambda_e,\delta)] \quad , \tag{45}$$

1.- Minimiser $-L(\omega^{**})$ est équivalent à maximiser $L(\omega^{**}) = -L(\omega^{**})$.

2.- Nous éliminons les indices ℓ, k, m, pour ρ, λ_x, λ_e, δ afin d'indiquer les vecteurs correspondants.

et[3] $\tilde{\beta}(\rho, \lambda_y, \lambda_x, \lambda_e, \delta) = (X^{**'} X^{**})^{-1} X^{**'} y^{**}$. (46)

La concentration de la fonction sur σ^{**2} implique que les estimateurs β qui maximisent $L(\omega^{**})$ pour des valeurs données de ρ, λ_y, λ_x, λ_e, δ sont ceux des moindres carrés ordinaires. Malgré le fait que ces estimateurs sont calculés comme ceux des moindres carrés ordinaires, ils possèdent néanmoins les propriétés asymptotiques des estimateurs du maximum de vraisemblance (la convergence et l'efficacité asymptotique).

Afin de maximiser[4] la fonction (44) par rapport à ρ, λ_y, λ_x, λ_e, δ, nous utilisons dans la présente étude, l'algorithme de minimisation construit par Fletcher et Powell (1963). Cette méthode nécessite de définir analytiquement les dérivées premières de la fonction. Puisque cet algorithme en est un de minimisation de la fonction de vraisemblance, il s'agit tout simplement d'inverser le signe de la fonction et de ses dérivées.

Si la fonction est continue, la condition suffisante qui nous assure que nous avons atteint un maximum au moins local est que les dérivées premières soient égales à zéro. Cependant, puisque le gradient n'est pas invariant aux unités de mesure, il est possible que numériquement, le gradient ne s'annule pas et que le processus génère quand même une solution qui corresponde à un maximum au moins local de la fonction de vraisemblance.

Si nous sommes assurés que le maximum atteint est au moins local, rien

---

3.- Pour une discussion plus détaillée sur le calcul de $\tilde{\beta}$, $\tilde{\sigma}^{**2}$, le logarithme de la fonction de vraisemblance non concentrée et concentrée sur β et sur σ^{**2}, voir appendice 1.

4.- Un programme (DSDM: Demande Spatiale Directe par Mode) pour calculer la solution du maximum de vraisemblance impliquant un grand nombre de paramètres (v.g. équation 44) a été développé par Tran Cong Liem (1979). Ce programme nous permet soit de fixer, soit de laisser estimer les différents paramètres impliqués dans $L(\omega^{**})$. Sa très grande flexibilité en fait donc un outil de travail extrêmement riche.

nous indique cependant que ce maximum en soit un global. Pour nous assurer que le maximum est global, il faudrait vérifier si le point de convergence est robuste à la valeur initiale - c'est-à-dire vérifier si la fonction de vraisemblance converge au même point peu importe la solution de départ et ce, pour l'ensemble des solutions de départ raisonnables-[5]. S'il était démontré qu'il n'y a pas de solution unique, il faudrait alors trouver tous les maxima locaux et choisir le maximum maximorum.

Toutefois, il est impensable de chercher à voir si la fonction converge à un maximum global, compte tenu du coût monétaire et du coût temps qu'entraînerait la localisation d'un maximum global.

Si la convergence numérique est assurée et que de ce fait un maximum (au moins local) de $L(\omega^{**})$ est atteint, nous pouvons alors calculer la matrice de covariance asymptotique pour les valeurs finales des paramètres[6] ρ, λ_y, λ_x, λ_e, δ (obtenues à la dernière itération lors du processus de minimisation de $-L(\omega^{**})$) qui maximisent la fonction de vraisemblance et pour les β.

La matrice de covariance pour les paramètres précités peut être estimée en utilisant la relation statistique que "la matrice de covariance asymptotique d'un estimateur du maximum de vraisemblance est égale à l'inverse de la matrice de covariance du gradient de la fonction de vraisemblance", (Kendall et Stuart, 1967, p.55).

Le gradient de la fonction de vraisemblance logarithmique concentrée

---

5.- Pour maximiser sur l'espace entier des paramètres, il faudrait choisir des valeurs alternatives de λ sur un espace raisonnable et régresser y sur X. Pour éviter des problèmes numériques, la région permise est $|\lambda| < 20$.

6.- Dans le présent rapport, la covariance asymptotique entre σ^{**2} et les autres paramètres est posée comme étant nulle, de sorte que l'expression (48) telle que définie dans la page 29, n'est pas tout à fait exacte.
Par conséquent, la procédure pour le calcul des écarts-types des divers paramètres est approximative et inexacte.

sur σ^{**2} peut être écrit comme étant:

$$\frac{\partial L(\omega^*)}{\partial(\omega^*)} = \sum_{t=1+r}^{n} D_t \qquad (47)$$

où (ω^*) est un vecteur colonne de M éléments constitué par les paramètres β,

ρ, λ_y, λ_x, λ_e, δ,

et

où D_t est un vecteur colonne de M éléments[7];

$$D_t = \frac{\partial (\lambda_y - 1) \ln y_t - (y_t^{**} - \sum_k \beta_k X_{k_t}^{**}) \; \partial (y_t^{**} - \sum_k \beta_k X_{k_t}^{**})}{\partial(\omega^*)} \qquad \hat{\sigma}^{**2} \qquad \partial(\omega^*)$$

$$- \frac{1}{2} \frac{\partial \ln f(X_t)}{\partial(\omega^*)} . \qquad (48)$$

Au maximum de la fonction de vraisemblance, la matrice de covariance asymptotique d'ordre M × M pour $\omega^* = \hat{\omega}^*$ est donnée par:

$$\text{COV}(\hat{\omega}^*) = \left[\sum_{t=1+r}^{n} D_t D'_t \right]^{-1}_{\omega^* = \hat{\omega}^*} , \qquad (49)$$

où les éléments paramétriques du vecteur $\hat{\omega}^*$ sont les estimés des éléments du

---

7.- Le second terme de D_t vient de:

$$\partial \left[- \frac{(n-r/2) \ln \hat{\sigma}^{**2}}{\partial(\omega^*)} \right] = - \left(\frac{n-r}{2} \right) \frac{1}{\hat{\sigma}^{**2}} \cdot \frac{\partial \hat{\sigma}^{**2}}{\partial(\omega^*)} \qquad (1)$$

où $\hat{\sigma}^{**2} = \frac{1}{(n-r)} e'_t e_t$ (annexe 1, équation 1.5)(2), avec $e_t = y_t^{**} - \sum_k \beta_k X_{k_t}^{**}$.

Si nous substituons (2) dans (1), nous avons:

$$- \left(\frac{n-r}{2} \right) \frac{1}{\hat{\sigma}^{**2}} \cdot \frac{1}{(n-r)} 2 \left[\frac{\partial e'_t}{\partial(\omega^*)} \cdot e_t \right] = - \frac{e_t}{\hat{\sigma}^{**2}} \cdot \frac{\partial e'_t}{\partial(\omega^*)}$$

$$= - \frac{(y_t^{**} - \sum_k \beta_k X_{k_t}^{**}) \cdot \partial(y_t^{**} - \sum_k \beta_k X_{k_t}^{**})}{\hat{\sigma}^{**2} \qquad \partial(\omega^*)}$$

vecteur ω^*.

Les différentes dérivées partielles impliquées dans D_t sont présentées dans le texte de Tran Cong Liem (1979, pp.19-22).

Finalement, nous pouvons ajouter que les estimés de β et de σ^{**2} peuvent être obtenus des résultats de régression à $\hat{\lambda}$, $\hat{\rho}$, $\hat{\delta}$.

Dans le présent rapport de recherche, nous allons maximiser $L(\omega^{**})$ compte tenu de diverses hypothèses que nous émettrons à priori sur les différents paramètres inclus dans (ω^{**}). Dans un chapitre subséquent, nous mettons en exergue les diverses hypothèses émises (restrictives ou non).

* * *

Présentons maintenant le modèle général, lequel est obtenu suite à une combinaison des équations (27), (28) et (29'). Ainsi,

$$\frac{y_t^{(\lambda_y)}}{\sqrt{\exp(\sum_m \delta_m X_{m_t}^{(\lambda_{em})})}} - \sum_{\ell=1}^{r} \rho_\ell \frac{y_{t-\ell}^{(\lambda_y)}}{\sqrt{\exp(\sum_m \delta_m X_{m_{t-\ell}}^{(\lambda_{em})})}} =$$

$$\sum_k \beta_k \frac{X_{k_t}^{(\lambda_{xk})}}{\sqrt{\exp(\sum_m \delta_m X_{m_t}^{(\lambda_{em})})}} - \sum_k \sum_{\ell=1}^{r} \beta_k \rho_\ell \frac{X_{k_{t-\ell}}^{(\lambda_{xk})}}{\sqrt{\exp(\sum_m \delta_m X_{m_{t-\ell}}^{(\lambda_{em})})}} + e_t \qquad (50)$$

avec $t=1+r, \ldots, n$ et $\ell=1, \ldots, r$,

et où e_t respecte les mêmes hypothèses[8] que s_t.

8.- En ce qui concerne l'hypothèse de normalité de l'erreur, voir appendice 2.

Dans les modèles en séries chronologiques, l'identification d'un schéma autorégressif pertinent peut poser un problème. Pour estimer les paramètres du modèle de l'équation (50), nous identifions une structure autorégressive dans l'esprit des techniques de Box et Jenkins (1970, pp.496-497) en examinant les fonctions d'autocorrélation et d'autocorrélation partielle des estimateurs non-biaisés des résidus[9] Ω_t, lesquels sont obtenus en appliquant les moindres carrés ordinaires[10] à l'équation (27) pour $\lambda_y = \lambda_{xk} = 1$ (Gaudry et Wills, 1978, p.53). Nous obtenons alors une spécification initiale de la structure autorégressive, laquelle spécification est considérée par la suite dans l'estimation conjointe de la forme fonctionnelle et de la partie stochastique du modèle. Par conséquent, la structure d'autocorrélation est établie indépendamment de la forme fonctionnelle de la partie fixe du modèle même si nous savons implicitement qu'elle en est dépendante. Ceci est dûe à ce que nous sommes intéressés entre autres par l'impact des paramètres λ_y, λ_{xk}, λ_{em}, δ_m sur les estimateurs d'une structure donnée de paramètres β_k et par conséquent, il serait inapproprié d'apporter continuellement des changements au schéma autorégressif.

Par la suite, les fonctions d'autocorrélation et d'autocorrélation partielle sont évaluées pour les résidus e_t à la fin du processus de maximisation de $L(\omega^{**})$ et il est alors possible de vérifier si le schéma autorégressif est bien défini pour chacune des fonctions à maximiser.

9.- Si $\lambda_{em} = \delta_m = 0$, nous avons $u_t = v_t = \rho_\ell \, v_{t-\ell} + e_t$.

10.- Ici nous faisons allusion à un pseudo-moindres carrés ordinaires puisqu'une constante égale à l'unité est soustraite des valeurs observées $Z^{(\lambda=1)}$ où $(Z^{(\lambda=1)} - 1/\lambda=1)$.

CHAPITRE III

MODELE A ESTIMER

3.1. Application du modèle général à une fonction de demande d'essence.

Le cadre analytique d'une théorie qui prend en considération la détermination simultanée de la forme fonctionnelle de la partie fixe du modèle et de la spécification générale de l'hétéroscédasticité et qui généralise la procédure Box-Cox aux erreurs reliées temporellement est maintenant bien établi tout comme la procédure d'estimation.

Dès lors, nous pouvons, à cette étape de notre étude, présenter le modèle d'estimation considéré.

Le contexte opérationnel choisi est assujetti à une fonction de demande d'essence ou encore à une fonction de demande de transport en automobile dans une zone spatiale déterminée, à savoir la province de Québec.

Les diverses composantes qui caractérisent directement la fonction de demande d'essence ou de demande pour un mode de transport spécifique se réfèrent à des facteurs socio-économiques, à des facteurs reliés au transport, à des variables d'activité et autres attributs.

Nous pouvons représenter la fonction de demande d'essence telle que développée par M. Gaudry (1978), sous la forme algébrique suivante:

$$y_t^d = f^d (P_t, I_t, T_t, C_t, A_t, O_t) \qquad (51)$$

où

y_t^d : représente la quantité consommée agrégée d'essence à moteur,

P_t : est un vecteur contenant les variables prix: le prix du mode

de transport considéré et celui du mode concurrent, le coût réel de l'essence et l'indice général des prix,

I_t : représente le revenu,

T_t : est un vecteur de variables de niveau de service (variables de temps),

C_t : est un vecteur de variables de confort (variables de température),

A_t : est un vecteur de niveaux d'activité,

O_t : est un vecteur d'attributs divers.

La demande d'essence est analysée à partir de séries temporelles mensuelles consécutives correspondant à la période qui s'étend de décembre 1956 à juin 1977 inclusivement. L'échantillon est alors composé de 247 observations.

3.1.1. Paramétrisation de la partie fixe du modèle.

La transformation paramétrique Box-Cox est applicable, tel que nous l'avons noté antérieurement,à des variables positives. Or, certaines des 22 variables explicatives inscrites dans la liste des régresseurs ne respectent pas cette condition et ne sont par conséquent sujettes à la paramétrisation.

Dans la spécification de notre modèle, nous pouvons discerner deux groupes distincts de variables.

- Le premier groupe s'identifie aux variables pour lesquelles les données du vecteur correspondant sont:
  strictement positives: $x_t > 0$.
- Le second groupe est constitué par des variables "semi-dichotomiques", c'est-à-dire par des variables pour lesquelles les données du vecteur correspondant sont:
  soit positives : $x_i > 0$, i=nombre d'observations positives
  soit nulles : $x_j = 0$, j=nombre d'observations nulles

où $i + j = t$.

A partir du tableau 1, nous constatons que toutes les composantes des vecteurs P_t, T_t, O_t et du scalaire I_t peuvent être transformées. Cependant, 5 des 7 variables d'activité et les différentes variables de confort ne peuvent subir de transformations Box-Cox.

Le programme DSDM permet de traiter simultanément jusqu'à 6 transformations distinctes de la variable dépendante et de groupes de variables indépendantes[1]. Aussi compte tenu du nombre élevé de variables explicatives dans le modèle à l'étude, il nous faut regrouper les variables homogènes en un groupe distinct et les contraindre à avoir un lambda commun. Les variables du groupe confort ne pouvant subir de transformations, les cinq groupes restants (tels que définis dans (51)) peuvent alors conserver leur caractère homogène.

La structure de la paramétrisation de la partie fixe du modèle étant maintenant bien établie, nous pouvons représenter l'équation (51) sous une forme Box-Cox, laquelle spécification est un cas spécial de la forme générale présentée dans (27), (28) et (29'). Le modèle peut être écrit comme étant:

$$y_t^{(\lambda_y)} = \sum_h \sum_{k=1}^{4} \beta_{hk} P_{hk_t}^{\lambda_{x(1)}} + \beta_5 I_t^{\lambda_{x(2)}} + \sum_h \sum_{k=6}^{8} \beta_{hk} T_{hk_t}^{\lambda_{x(3)}}$$

$$+ \sum_{k=9}^{12} \beta_k C_{k_t} + \left(\sum_{k=15,16} \beta_k A_{k_t}^{\lambda_{x(4)}} + \sum_{k=13,14,17}^{19} \beta_k A_{k_t} \right)$$

$$+ \sum_{k=20}^{22} \beta_k O_{k_t}^{\lambda_{x(5)}} + u_t \qquad (52)$$

où (1) ,....., (5)=g et où g représente le g ième groupe de X_k et,

---

1.- Le programme permet également d'associer douze paramètres à la forme de l'hétéroscédasticité des erreurs résiduelles et cinq à la structure d'autocorrélation.

TABLEAU 1

Nombre d'observations positives pour chacun des régresseurs.

Groupe	Code	Variable[2]	Nbr. obs. positives
Prix (P)	1. FADCB	Tarif réel du transport en commun	247
	2. PGRMD	Coût réel de l'essence	247
	3. NCQR	Nombre d'automobiles	247
	4. PCB	Indice des prix à la consommation	247
Revenu (I)	5. RAWEQ	Salaire réel moyen	247
Temps (T)	6. TW	Temps d'attente du t.e.c.	247
	7. TT	Temps de transport en t.e.c.	247
	8. TAUTO	Temps de transport en automobile	247
Confort (C)	9. TLK	Inconfort dû à la température	246
	10. RFK	Inconfort dû à la pluie	244
	11. SFK	Inconfort dû à la neige	133
	12. CSFMR	Inconfort dû à la neige accumulée	131
Activités (A)	13. RCUTPN	Indice de présences scolaires totales aux niveaux primaire et secondaire	242
	14. ISECON	Indice de présences scolaires aux collèges et universités	207
	15. RSIDCB	Indice de déplacements-magasinage	247
	16. EIQP	Niveau de l'emploi	247
	17. EX7	Expo'67	7
	18. STR	Grève à la C.T.C.U.M.	9
	19. VWM	Congé des travailleurs	40
Agrégation (O)	20. WD	Jours de travail/mois	247
	21. SAT	Samedis/mois	247
	22. SHD	Dimanches & jours de fêtes/mois	247

2.- Pour la définition des variables, voir appendice 4.

où k=1 ,...., 22 est un indice relié à chacune des variables,

h=1 , 2 est un indice inféodé au mode,

avec $u_t = \sqrt{f(X_t)}\, v_t$

$$v_t = \sum_{\ell=1}^{r} \rho_\ell\, v_{t-\ell} + e_t \ .$$

L'ordre d'autocorrélation le plus élevé est le 12°. Les 1° et 2° ordres sont également considérés. Alors,

- l'indice r=12
- l'indice t=13 ,...., 247.

3.1.2. Paramétrisation de la partie aléatoire du modèle.

Dans ce genre de modèle en séries chronologiques, nous pouvons penser (comme nous le suggère la théorie traditionnelle) que les variables les plus susceptibles d'influencer dans le temps le nombre de déplacements mensuels ou encore la demande pour un bien complémentaire à un mode de transport sont les variables qui sont affectées par des mouvements saisonniers. Aussi les variables d'activité, indice de déplacements pour motif de magasinage (*RSIDCB*) et niveau de l'emploi (*EIQP*) sont celles qui parmi toutes les variables strictement positives sont les plus susceptibles d'engendrer de l'hétéroscédasticité de la variable aléatoire. Cependant comme nous travaillons dans le champ de la non-linéarité, nous ne savons pas à priori quel sera leur comportement.

Nous retiendrons deux formes d'hétéroscédasticité des erreurs résiduelles.

1°) La formulation générale de l'hétéroscédasticité, lorsque $\lambda_u = 0$, à savoir:

$$f(X_t) = \exp\left(\sum_m \delta_m X_{m_t}^{(\lambda_{em})}\right)$$

2°) La formulation classique de l'hétéroscédasticité, laquelle est un cas spécial de la forme générale lorsque $\lambda_{em} = 0$ et $\delta_m = 2$ à savoir:

$$f(X_t) = \prod_m X_{m_t}^{\delta_m} \quad .$$

Ces deux formes d'hétéroscédasticité seront d'abord testées pour chacune des deux variables choisies pour X_{m_t} prises séparément et puis pour les deux prises conjointement.

Dans le cas où nous testons pour la forme générale avec les deux variables prises conjointement, nous estimons deux λ_{em} distincts, de même que deux δ_m. (Le programme ne permet pas d'imposer sur ces paramètres une contrainte d'égalité.)

3.2. Nomenclature des hypothèses.

La transformation Box-Cox sur les parties fixe et stochastique du modèle et la très grande flexibilité du programme utilisé, nous permettent d'établir une liste très exhaustive sur les différentes possibilités d'hypothèses que nous pouvons formuler.

Dans le but d'obtenir des résultats générateurs d'indices sur la capacité et sur l'utilité du modèle général à nous informer sur le discernement possible (contrairement aux procédures existantes) entre la compensation ou la non-compensation de l'hétéroscédasticité des résidus par les paramètres de la forme fonctionnelle, sur la linéarité ou la non-linéarité[3] réellement présente

---

3.- Une des idées à la base de cette procédure est de pouvoir discerner de façon concomitante entre:
- la mesure dans laquelle les paramètres λ estimés de la forme fonctionnelle compensent pour l'hétéroscédasticité des résidus,
- la mesure dans laquelle ils reflètent la non-linéarité réellement présente dans le modèle (Gaudry, 1978, p.17).

dans le modèle, ainsi que sur son utilité à nous informer sur les effets de la forme fonctionnelle "optimale" laquelle est inhérente à la détermination conjointe de la forme fonctionnelle des parties fixe et stochastique, nous allons déterminer différentes combinaisons de transformation sur les parties fixe et aléatoire du modèle.

3.2.1. Hypothèses sur la partie fixe.

Nous émettons six séries d'hypothèses sur la partie fixe du modèle. Dans chacune de ces séries, les variables positives sont paramétrisées. Dans la première, les paramètres sont contraints à être fixés et dans les cinq autres, ils sont estimés.

Série $\star$: Nous appelons série $\star$, lorsque nous contraignons tous les éléments du vecteur λ (le lambda y sur la variable dépendante et le lambda x sur chacun des cinq groupes (prix, revenu, temps, activités et autres)) à être fixés et à prendre une forme soit linéaire, soit logarithmique.

Série A: Dans une seconde série (A), nous estimons les différents paramètres en les contraignant tous à être égaux entre eux.

Série B: Dans la série B, nous libérons la contrainte d'égalité entre le lambda y et les lambdas x, tout en conservant la contrainte d'égalité entre ces derniers.

Dans chacune des trois séries présentées ci-dessous, il y a également la relaxation de la contrainte d'égalité entre le lambda y et les lambdas x.

Série C: Nous postulons de plus dans cette série que les paramètres affectés sur les variables incluses dans les groupes P_t, I_t, T_t, O_t seront estimés en étant contraints à être égaux, cependant que le lambda x sur les variables incluses dans le groupe A_t est libéré de toutes

contraintes d'égalité.

Série D: Dans la série D, nous retrouvons les mêmes restrictions que celles émises dans C sauf que le lambda x sur les variables incluses dans le groupe O_t est également libéré de toutes contraintes d'égalité. Ainsi les lambdas x affectant les groupes A_t et O_t sont estimés librement tandis que les lambdas x des groupes P_t, I_t et T_t sont estimés en étant contraints à être égaux entre eux.

Série E: Un relâchement complet de toutes restrictions d'égalité entre les différents lambdas nous permet d'obtenir le modèle le moins conditionnel donc le plus général.

La relaxation des égalités va nous permettre d'analyser le comportement du lambda y et des lambdas x.

3.2.2. Hypothèses sur la partie aléatoire.

Nous formulons également six hypothèses distinctes sur les paramètres de la partie stochastique.

Dans la notation utilisée pour la partie aléatoire du modèle, nous retrouvons les symboles suivants:

HG: hétéroscédasticité générale

HC: hétéroscédasticité classique

AU: autocorrélation.

Hypothèse 2: BC_S

Le symbole BC est utilisé pour dénoter la transformation Box-Cox. L'indice S est relié aux différentes séries. Ainsi pour chacune des hypothèses formulées à priori sur la partie fixe, nous associons différentes hypothèses sur la partie aléatoire. Cette hypothèse définit un modèle où tous les

paramètres de la partie aléatoire (les éléments du vecteur ρ, du vecteur λ_e et du vecteur δ) sont contraints à être égaux à zéro.

Hypothèse 3: $BC_S AU$

Dans cette hypothèse, nous postulons la présence d'un schéma autorégressif des erreurs résiduelles. Par conséquent, nous ne contraignons plus dans ce modèle les éléments du vecteur des ρ (ρ_1 , ρ_2, ρ_{12}) à être égaux entre eux et à être égaux à zéro comme dans l'hypothèse précédente.

Hypothèse 4: $BC_S HG_\Phi$

Cette hypothèse établit un modèle dans lequel nous introduisons la forme générale de l'hétéroscédasticité telle que définie précédemment.
Si la variable *RSIDCB* est choisie pour X_m, alors nous posons $\Phi = 1$,
Si la variable *EIQP* est choisie pour X_m, alors nous posons $\Phi = 2$,
Si les 2 variables sont traitées simultanément, nous avons $\Phi = 1, 2$.

Les paramètres associés au(x) source(s) d'hétéroscédasticité sont estimés.

Hypothèse 5: $BC_S HC_\Phi$

Le modèle formulé dans cette hypothèse est un cas particulier de celui formulé précédemment. Il introduit une forme d'hétéroscédasticité classique. Les paramètres de la partie stochastique sont posés comme étant fixés à une valeur particulière.

Hypothèse 6: $BC_S AUHG_\Phi$

Parmi toutes les hypothèses, celle-ci est la plus générale ou la moins contraignante. Contrairement à l'hypothèse 4, la transformation Box-Cox sur la partie aléatoire du modèle est appliquée aux erreurs qui sont temporellement corrélées.

Hypothèse 7: $BC_S AUHC_\Phi$

Nous ajoutons à l'hypothèse 5, une structure d'autocorrélation des erreurs, de sorte que la forme de l'hétéroscédasticité classique est analysée conjointement avec les erreurs autorégressives.

Dans le tableau 2, nous résumons les différents modèles formulés.

Notons qu'un indice double **, AA, BB, CC, DD, EE y dénote l'utilisation de deux variables d'hétéroscédasticité.

TABLEAU 2

Tableau récapitulatif sur les combinaisons d'hypothèses sur les parties fixe et aléatoire.

Test	Modèle	Hypothèses sur la partie fixe	Hypothèses sur la partie aléatoire
2*	BC_*	$\lambda_y=\lambda_{x1}=\lambda_{x2}=\lambda_{x3}=\lambda_{x4}=\lambda_{x5}=1$ ou $=0$	$\rho_1=\rho_2=\rho_{12}=0;\lambda_{e1}=0,\lambda_{e2}=0;\delta_1=0,\delta_2=0$
3*	BC_*AU		$\hat{\rho}_1,\hat{\rho}_2,\hat{\rho}_{12}$, $;\lambda_{e1}=0,\lambda_{e2}=0;\delta_1=0,\delta_2=0$
4*	BC_*HG_1		$\rho_1=\rho_2=\rho_{12}=0;\hat{\lambda}_{e1}$ $,\lambda_{e2}=0;\hat{\delta}_1$ $,\delta_2=0$
	BC_*HG_2		$\rho_1=\rho_2=\rho_{12}=0;\lambda_{e1}=0,\hat{\lambda}_{e2}$ $;\delta_1=0,\hat{\delta}_2$
5*	BC_*HC_1		$\rho_1=\rho_2=\rho_{12}=0;\lambda_{e1}=0,\lambda_{e2}=0;\delta_1=2,\delta_2=0$
	BC_*HC_2		$\rho_1=\rho_2=\rho_{12}=0;\lambda_{e1}=0,\lambda_{e2}=0;\delta_1=0,\delta_2=2$
6*	BC_*AUHG_1		$\hat{\rho}_1,\hat{\rho}_2,\hat{\rho}_{12}$ $;\hat{\lambda}_{e1}$ $,\lambda_{e2}=0;\hat{\delta}_1$ $,\delta_2=0$
	BC_*AUHG_2		$\hat{\rho}_1,\hat{\rho}_2,\hat{\rho}_{12}$ $;\lambda_{e1}=0,\hat{\lambda}_{e2}$ $;\delta_1=0,\hat{\delta}_2$
7*	BC_*AUHC_1		$\hat{\rho}_1,\hat{\rho}_2,\hat{\rho}_{12}$ $;\lambda_{e1}=0,\lambda_{e2}=0;\delta_1=2,\delta_2=0$
	BC_*AUHC_2		$\hat{\rho}_1,\hat{\rho}_2,\hat{\rho}_{12}$ $;\lambda_{e1}=0,\lambda_{e2}=0;\delta_1=0,\delta_2=2$
2A	BC_A	$\hat{\lambda}_y=\hat{\lambda}_{x1}=\hat{\lambda}_{x2}=\hat{\lambda}_{x3}=\hat{\lambda}_{x4}=\hat{\lambda}_{x5}$	idem

⋮			
7A	$BC_A AUHC_2$		idem
2B	BC_B	$\hat{\lambda}_y \neq \hat{\lambda}_{x1} = \hat{\lambda}_{x2} = \hat{\lambda}_{x3} = \hat{\lambda}_{x4} = \hat{\lambda}_{x5}$	idem
⋮			
7B	$BC_B AUHC_2$		idem
2C	BC_C	$\hat{\lambda}_y \neq \hat{\lambda}_{x1} = \hat{\lambda}_{x2} = \hat{\lambda}_{x3} = \hat{\lambda}_{x5} \neq \hat{\lambda}_{x4}$	idem
⋮			
7C	$BC_C AUHC_2$		idem
2D	BC_D	$\hat{\lambda}_y \neq \hat{\lambda}_{x1} = \hat{\lambda}_{x2} = \hat{\lambda}_{x3} \neq \hat{\lambda}_{x4} \neq \hat{\lambda}_{x5}$	idem
⋮			
7D	$BC_D AUHC_2$		idem
2E	BC_E	$\hat{\lambda}_y \neq \hat{\lambda}_{x1} \neq \hat{\lambda}_{x2} \neq \hat{\lambda}_{x3} \neq \hat{\lambda}_{x4} \neq \hat{\lambda}_{x5}$	idem
⋮			
7E	$BC_E AUHC_2$		idem
4**	$BC_{**} HG_{1,2}$	idem à série *	$\rho_1=\rho_2=\rho_{12}=0; \hat{\lambda}_{e1}, \hat{\lambda}_{e2}; \hat{\delta}_1, \hat{\delta}_2$
5**	$BC_{**} HC_{1,2}$		$\rho_1=\rho_2=\rho_{12}=0; \lambda_{e1}=0, \lambda_{e2}=0; \delta_1=2, \delta_2=2$
6**	$BC_{**} AUHG_{1,2}$		$\hat{\rho}_1, \hat{\rho}_2, \hat{\rho}_{12}; \hat{\lambda}_{e1}, \hat{\lambda}_{e2}; \hat{\delta}_1, \hat{\delta}_2$
7**	$BC_{**} AUHC_{1,2}$		$\hat{\rho}_1, \hat{\rho}_2, \hat{\rho}_{12}; \lambda_{e1}=0, \lambda_{e2}=0; \delta_1=2, \delta_2=2$
AA		idem à série A	idem à série **
BB		idem à série B	idem
CC		idem à série C	idem
DD		idem à série D	idem
EE		idem à série E	idem

3.3. Comparaison entre les hypothèses.

Parmi les hypothèses présentées précédemment, la logique suggère de comparer les valeurs maximales des fonctions de vraisemblance (critère fondamental) deux à deux à l'aide du test du rapport des vraisemblances pour les hypothèses suivantes.

. absence d'autocorrélation.

T. 1 $BC_{\star}$ VS $BC_{\star}HG_1$

$[-\hat{L}(\lambda_y=\lambda_{x1}=\lambda_{x2}=\lambda_{x3}=\lambda_{x4}=\lambda_{x5}=1 \text{ ou } 0;\rho_1=\rho_2=\rho_{12}=0;\lambda_{e1}=0,\lambda_{e2}=0;\delta_1=0,\delta_2=0)$

$+\hat{L}(\lambda_y=\lambda_{x1}=\lambda_{x2}=\lambda_{x3}=\lambda_{x4}=\lambda_{x5}=1 \text{ ou } 0;\rho_1=\rho_2=\rho_{12}=0;\hat{\lambda}_{e1},\lambda_{e2}=0;\hat{\delta}_1,\delta_2=0)]$

nous avons également T. 1 pour . $BC_{\star}$ VS $BC_{\star}HG_2$

. $BC_{\star}$ VS $BC_{\star\star}HG_{1,2}$.

Les restrictions pour chacun des modèles étant déjà spécifiées dans la section 3.2., nous poursuivrons en ne mentionnant que la codification de chacun d'eux.

T. 2 $BC_{\star}$ VS BC_A

T. 3 BC_A VS $BC_{\star}HG_1$

nous avons également T. 3 pour . BC_A VS $BC_{\star}HG_2$

. BC_A VS $BC_{\star\star}HG_{1,2}$

T. 4 BC_A VS BC_AHG_1

nous avons également T. 4 pour . BC_A VS BC_AHG_2

. BC_A VS $BC_{AA}HG_{1,2}$

T. 5 BC_A VS BC_AHC_1

nous avons également T. 5 pour . BC_A VS BC_AHC_2

. BC_A VS $BC_{AA}HC_{1,2}$

T. 6 BC_AHC_1 VS BC_AHG_1

nous avons également T. 6 pour . BC_AHC_2 VS BC_AHG_2

. $BC_{AA}HC_{1,2}$ VS $BC_{AA}HG_{1,2}$

T. 7 BC_AHG_1 VS $BC_{AA}HG_{1,2}$

nous avons également T. 7 pour . BC_AHG_2 VS $BC_{AA}HG_{1,2}$

T. 8 BC_AHC_1 VS $BC_{AA}HC_{1,2}$

nous avons également T. 8 pour . BC_AHC_2 VS $BC_{AA}HC_{1,2}$

. présence d'autocorrélation.

T. 9 BC_* VS BC_*AU

T. 9' BC_A VS BC_AAU

T.10 BC_*AU VS BC_*AUHG_1

nous avons également T.10 pour . BC_*AU VS BC_*AUHG_2

. BC_*AU VS $BC_{**}AUHG_{1,2}$

T.11 BC_*AU VS BC_AAU

T.12 BC_AAU VS BC_*AUHG_1

nous avons également T.12 pour . BC_AAU VS BC_*AUHG_2

. BC_AAU VS $BC_{**}AUHG_{1,2}$

T.13 BC_AAU VS BC_AAUHG_1

nous avons également T.13 pour . BC_AAU VS BC_AAUHG_2

. BC_AAU VS $BC_{AA}AUHG_{1,2}$

T.14 BC_AAU VS BC_AAUHC_1

nous avons également T.14 pour . BC_AAU VS BC_AAUHC_2

. $BC_A AU$ VS $BC_{AA} AUHC_{1,2}$

T.15 $BC_A AUHC_1$ VS $BC_A AUHG_1$

nous avons également T.15 pour . $BC_A AUHC_2$ VS $BC_A AUHG_2$

. $BC_{AA} AUHC_{1,2}$ VS $BC_{AA} AUHG_{1,2}$

T.16 $BC_A AUHG_1$ VS $BC_{AA} AUHG_{1,2}$

nous avons également T.16 pour . $BC_A AUHG_2$ VS $BC_{AA} AUHG_{1,2}$

T.17 $BC_A AUHC_1$ VS $BC_{AA} AUHC_{1,2}$

nous avons également T.17 pour . $BC_A AUHC_2$ VS $BC_{AA} AUHC_{1,2}$.

Les tests T.2, T.3, T.4, T.5, T.6, T.7, T.8, T.9', T.11, T.12, T.13, T.14, T.15, T.16, T.17 peuvent être également répétés pour toutes les séries que nous avons mentionnées.

Les tests T.1 et T.10 ont pour but de permettre la détection de la présence ou de l'absence de l'hétéroscédasticité dans un modèle dans lequel les paramètres de la partie fixe sont contraints à prendre une valeur particulière.

Les tests T.2 et T.11 vont nous permettre de voir si la forme fonctionnelle est modifiée d'une façon significative si on laisse estimer les paramètres de la partie fixe. Ces tests vont révéler l'influence de contraintes sur les λ, dans ce cas-ci à être égaux à 1.

L'intérêt suscité par les tests T.3 et T.12 est dû à ce que ces derniers nous permettent de voir si le paramètre lambda y de la partie fixe compense (en totalité ou en partie) ou ne compense pas pour l'hétéroscédasticité décelée par les tests T.1 et T.10.

Les tests T.4 et T.13 nous indiquent l'importance de la détermination

conjointe de la forme fonctionnelle de la partie fixe et de la partie stochastique du modèle.

Par essence, les tests T.5 et T.14 mesurent l'impact de la transformation Box-Cox sur la partie fixe si nous l'estimons conjointement avec la forme classique de l'hétéroscédasticité.

Les tests T.6 et T.15 vont nous instruire sur l'utilité de bien spécifier la formulation de l'hétéroscédasticité. Ils ont pour but d'évaluer les contraintes sur λ_{em} et δ_m: $\lambda_{em} = 0$, $\delta_m = 2$.

Les tests T.7 et T.16, ainsi que T.8 et T.17 permettent de vérifier si deux variables prises conjointement sont sources d'une plus grande présence de l'hétéroscédasticité.

Le test T.9 a pour but d'évaluer la contrainte placée sur les $\rho_\ell = \rho_1 = \rho_2 = \rho_{12} = 0$. Le test T.9' a pour but de juger de l'impact de la forme fonctionnelle sur l'autocorrélation.

La comparaison des modèles se fait dans les tests précités (sauf T.2 et T.11) sur les paramètres inclus dans le terme d'erreur, pour lesquels il y a des hypothèses conditionnelles et non-conditionnelles.

D'autres tests cependant peuvent être faits entre les différentes séries, lorsque nous maintenons les mêmes hypothèses sur le terme de l'erreur. Nous vérifions dans ce cas, l'impact sur la fonction de vraisemblance et sur les autres paramètres du modèle lorsque les hypothèses sur les paramètres de la partie fixe deviennent de moins en moins restrictives. Il s'agit alors de comparer $BC_A HG_1$ avec $BC_B HG_1$, BC_A avec BC_B, etc..

La statistique pertinente utilisée pour nous permettre de juger de l'importance de l'impact sur la fonction de vraisemblance pour chacun des tests

est, pour le processus de maximisation:

$$-\left[L_0(\omega^{**}) - L_1(\omega^{**})\right] = \boxed{\begin{array}{l}\left[-L_0(\omega^{**}) + L_1(\omega^{**})\right] \\ > 1/2\ \chi^2_\alpha\ (D.L.)\end{array}} \quad . \qquad (53)$$

Ainsi $L_0(\omega^{**})$ dénote la valeur logarithmique du maximum de la fonction de vraisemblance associée au modèle le plus contraint entre les deux modèles impliqués et $L_1(\omega^{**})$ dénote la valeur logarithmique du maximum de la fonction de vraisemblance associée au modèle le moins restrictif. Pour le test T.6 par exemple, $L_1(\omega^{**})$ est la valeur maximale de la fonction de vraisemblance lorsque $f(X_t)$ a une forme générale (expression 15) et $L_0(\omega^{**})$ sa valeur lorsque $f(X_t)$ a une forme particulière, telle que définie dans les expressions 18 ou 20. χ^2_α (D.L.) indique que la relation statistique (53) a une distribution asymptotique χ^2 pour certains degrés de liberté et un niveau de signification α.

Le nombre de degrés de liberté est égal à la différence dans le nombre de paramètres impliqués dans le processus de maximisation entre les deux modèles[4].

4.- La comparaison entre certains modèles nécessite l'utilisation de la distribution chi-carré avec un nombre de degrés de liberté moindre que la différence dans le nombre de paramètres.

. Par exemple, les tests T.1, T.4, T.10 et T.13 (lorsqu'une seule variable est impliquée dans $f(X_t)$)ont un seul degré de liberté même si la différence dans le nombre de paramètres entre les deux modèles impliqués s'élève à deux.

Prenons le test T.4.

Avec le modèle BC_AHG_1, nous avons $\hat{\delta}_1$, $\hat{\lambda}_{e1}$.

Avec le modèle BC_A , nous avons $\delta_1=0$, alors λ_{e1} peut prendre n'importe quelle valeur entre -20 et +20, si on le laisse estimer. (Nous le fixons égal à zéro, afin de ne pas l'introduire dans le processus de minimisation.)

Ainsi au point $\delta_1=0$, la dimension λ_{e1} n'existe pas et le test ne se fait que sur le delta.

Les points de pourcentage d'une distribution $1/2\chi^2$ pour 1 à 5 degrés de liberté sont les suivants:

Niveau / d.l.	0.10	0.05	0.025	0.01	0.005	0.001	0.0005	0.0001
1	1.36	1.92	2.51	3.32	3.94	5.41	6.05	7.57
2	2.31	3.00	3.69	4.61	5.30	6.91	7.60	9.21
3	3.13	3.91	4.68	5.67	6.42	8.15	8.85	
4	3.89	4.75	5.57	6.64	7.43	9.25	10.00	11.76
5	4.62	5.54	6.42	7.55	8.38	10.25	11.05	

. Les tests T.5, T.8, T.14 et T.17 soulèvent encore un point particulier. En effet le nombre de paramètres estimés est égal, pour les deux modèles impliqués.

Prenons le test T.5.

Pour le modèle BC_AHC_1, nous avons $\delta_1=2$, $\lambda_{e1}=0$.

Pour le modèle BC_A , nous avons $\delta_1=0$, λ_{e1} peut donc être estimé et prendre n'importe quelle valeur. Toutefois, puisque δ_1 multiplie $X_m{}^{\lambda_{e1}}$, nous avons le paramètre λ_{e1} qui s'estompe.

De plus les hypothèses sur δ_1 ($\delta_1=0$,$\delta_1=2$) ne sont pas emboîtées l'une dans l'autre. (Le test χ^2 suppose des modèles emboîtés l'un dans l'autre ("Nested hypothesis").)

A notre connaissance, il n'existe aucun test statistique qui puisse permettre la comparaison entre deux modèles lorsqu'il n'y a aucun degré de liberté.

Le seul critère de comparaison entre les modèles BC_AHC_1 et BC_A devient la valeur de la fonction de vraisemblance.

. Pour les tests T.3 et T.12, nous rencontrons le problème précédent (l'absence de degrés de liberté) si nous comparons BC_A à $BC_\star HG_1$ et BC_AAU à $BC_\star AUHG_1$. Ainsi pour

- BC_AAU , nous avons $\hat{\lambda}_y=\hat{\lambda}_{x1}=\hat{\lambda}_{x2}=\hat{\lambda}_{x3}=\hat{\lambda}_{x4}=\hat{\lambda}_{x5};\hat{\rho}_1, \hat{\rho}_2, \hat{\rho}_{12};\delta_1=0, \hat{\lambda}_{e1}$ (peut prendre n'importe quelle valeur).
- $BC_\star AUHG_1$, nous avons $\hat{\lambda}_y=\hat{\lambda}_{x1}=\hat{\lambda}_{x2}=\hat{\lambda}_{x3}=\hat{\lambda}_{x4}=\hat{\lambda}_{x5}=1;\hat{\rho}_1, \hat{\rho}_2, \hat{\rho}_{12};\hat{\delta}_1, \hat{\lambda}_{e1}$.

Toutefois si nous comparons BC_S à $BC_\star HG_1$ et BC_SAU à $BC_\star AUHG_1$, pour S différent de A, le problème ne se pose pas.

CHAPITRE IV

ANALYSE DES RESULTATS.

4.1. Etude sur le degré d'opérationnalité du modèle.

Le but visé dans la première sous-section de ce chapitre est de voir dans quelle mesure il est possible d'obtenir des résultats qui vont nous permettre de voir si les objectifs inféodés à la détermination conjointe de la forme fonctionnelle des parties fixe et stochastique sont réalisables.

Nous pouvons tout d'abord mentionner que le modèle de demande d'essence dans la province de Québec nous permet de justifier la rationalité de la pratique employée ici, même si nous avons rencontré quelques difficultés d'estimation.

La nature des problèmes est assez diversifiée.

Le problème le plus fréquemment rencontré est celui de la non-inversion de la matrice de covariance du gradient des paramètres. Lorsque ce problème se pose, nous obtenons seulement les valeurs des paramètres[1] impliqués dans le processus de maximisation et par conséquent la valeur finale de la fonction de vraisemblance.

1.- Durant le processus de maximisation pour $L(\omega^{**})$, les paramètres ρ_ℓ sont exprimés en termes de la transformation z de Fisher, où

$$z_\ell = \frac{1}{2} \ell n \frac{1 + \rho_\ell}{1 - \rho_\ell} \quad .$$

D'un problème de maximisation avec contraintes ($-1 < \rho_\ell < 1$) nous passons alors à un problème de maximisation sans contraintes ($-\infty < z_\ell < +\infty$).

La transformation inverse de z_ℓ est ρ_ℓ = tangente hyperbolique z_ℓ.

Il est intéressant de noter que l'introduction de ce problème est, dans de nombreux cas, reliée à une plus grande généralité des paramètres de la partie fixe. C'est ainsi que pour les modèles dans les séries A et B, ainsi que dans les séries AA et BB, le problème n'est pas apparu, contrairement aux expériences tentées dans les autres séries.

Règle générale, l'explication sous-jacente à cet aspect contrariant de la régression demeure (du moins à ce jour) indéterminée.

Ce problème peut également se poser si l'écart entre la valeur finale de la fonction de vraisemblance et la valeur de départ est trop faible. Il s'agit alors de modifier la situation initiale.

Le second problème que nous avons rencontré est celui où certains des paramètres impliqués dans le processus de maximisation atteignent une des bornes de la région admissible $|\lambda| < 20$. Deux des modèles probatoires de cette difficulté sont inféodés aux modèles $BC_{AA}HG_{1,2}$[2] et $BC_{BB}HG_{1,2}$. (Le paramètre λ_{e2} a atteint la borne supérieure dans chacun des deux modèles.) Dans de tels cas, la première propriété que se doit de posséder l'algorithme, à savoir la convergence numérique, n'est pas respectée (ou du moins, pas dans la région permise). Par conséquent, le processus ne pourra générer une solution qui corresponde à un maximum au moins local de la fonction de vraisemblance.

D'autre part, comme nous l'avons mentionné dans la seconde partie de ce rapport, il serait intéressant que le processus soit robuste aux modifications de la solution initiale pour pouvoir générer la propriété de convergence

---

2.- Dans le modèle $BC_{AA}HG_{1,2}$, λ_{e2} a atteint la borne de +20 et a par la suite régressé. Toutefois, même après un temps d'exécution de 5,000 secondes aucun point de convergence n'avait été atteint, et la valeur du paramètre λ_{e2} était approximativement 12. De plus, la fonction de vraisemblance était extrêmement plate.

vers l'unicité de la solution. Nous avons essayé pour certains de nos modèles, différentes solutions initiales et nos résultats démontrent d'une façon convaincante que cette propriété de robustesse semble assez bien respectée dans notre cas. Cependant, puisque nous n'avons pas "balayé" tout l'espace entre la région permise nous ne sommes pas assurés s'il s'agit de maxima globaux.

Finalement, le dernier problème soulevé, repose sur le non-respect du précepte qui suggère que la valeur de la fonction de vraisemblance doit s'accroître lorsque nous généralisons la procédure Box-Cox. Nous sommes confrontés à ce sérieux problème, lorsque nous estimons les modèles $BC_{**}AUHG_{1,2}$ et $BC_{AA}AUHG_{1,2}$. Dans le premier cas, la valeur finale de la fonction de vraisemblance s'élève à 179.64 et dans le second, elle s'élève à 179.06.

Pour les modèles dans lesquels aucun problème ne s'est posé, il apparaît que la fonction de vraisemblance de chacun d'eux se comporte assez bien dans l'espace des paramètres, puisque la procédure d'itérations converge assez rapidement si l'on considère que le nombre de paramètres qui maximisent la fonction de vraisemblance ainsi que le nombre d'observations est assez élevé.

Dans le tableau 3, nous indiquons le nombre d'itérations, ainsi que les temps et les coûts de résolution de quelques-uns de nos modèles.

TABLEAU 3

Nombre d'itérations, temps et coût de résolution de divers modèles.

Modèle	Nombre de paramètres	Nombre d'itérations	Temps (secondes)	Coût (dollars)
BC_AHG_1	3	27	588.587	66.215
BC_AHG_2	3	18	452.806	50.940
BC_CAUHG_2	8	27	1,501.703	168.941
BC_DAUHG_2	9	41	2,285.561	257.125

$BC_{AA}HC_{1,2}$	1	4	255.933	28.792
$BC_{AA}AUHG_{1,2}$	8	31	1,333.696	150.400
$BC_{EE}AUHG_{1,2}$	13	52	2,286.286	257.207

Dans les pages suivantes, nous allons présenter plusieurs des résultats que nous avons obtenus. Nous ne pouvons présenter la totalité de nos résultats et par conséquent, nous allons présenter dans la mesure du possible ceux qui, à notre avis, dénotent le mieux l'importance de la forme fonctionnelle et de la spécification de l'erreur.

4.2. Estimation et tests de la forme fonctionnelle.

4.2.1. Justification de l'utilité de la détermination conjointe, à partir d'une analyse des différents modèles sur la base des hypothèses émises sur la partie stochastique du modèle.

.Dans la présente sous-section, nous allons disserter sur un des objectifs fondamentaux de notre recherche. Explicitement, nous allons voir que la transformation paramétrique Box-Cox sur la partie fixe prise conjointement avec celle sur la partie stochastique détermine la forme fonctionnelle optimale du modèle à estimer. Les valeurs des paramètres lambdas de la partie fixe ne sont pas indépendantes des valeurs des paramètres qui spécifient la distribution des résidus, c'est-à-dire de celles des paramètres ρ_{ℓ}, λ_{em}, δ_{m} inclus dans (ω^{**}) et de ce fait, la détermination simultanée trouve la véritable essence de son fondement théorique. Par conséquent et contrairement aux pratiques courantes, il sera possible de voir dans quelle mesure le lambda estimé de la partie fixe est biaisé par son rôle compensateur ou stabilisateur si on ne corrige pas pour l'hétéroscédasticité alors que sa présence a été décelée.

L'application des tests sous-jacents aux modèles sans autocorrélation à savoir T.1, T.2, T.3 et T.4 et aux modèles avec autocorrélation à savoir T.9, T.10, T.11, T.12 et T.13 va nous permettre d'apporter un discernement discur-

sif entre la "vraie" forme fonctionnelle (ou du moins un modèle qui nous donne un meilleur ajustement)et la forme fonctionnelle "faussée" de par son rôle de compensation.

a) Deux variables prises séparément.

En premier lieu, nous allons analyser les implications sur la fonction de vraisemblance et sur la forme fonctionnelle si nous introduisons la variable *RSIDCB* comme source possible de l'hétéroscédasticité.

Soit le tableau 4 dans lequel sont inscrites les valeurs des différents paramètres et les valeurs correspondantes des fonctions de vraisemblance.

Cas 1

a) L'application du test T.1 pour le modèle BC_* versus le modèle BC_*HG_1, nous permet de déceler la présence de l'hétéroscédasticité des erreurs à un niveau de confiance de seulement 90% et de rejeter celle-ci avec confiance à un niveau de 95%. En effet, la valeur du test statistique pour une comparaison de cette paire d'hypothèses est égale à 1.73.

b) Puisque nous avons décelé une faible présence de l'hétéroscédasticité des erreurs résiduelles (si nous considérons la première alternative) nous allons, en laissant estimer le paramètre de la partie fixe du modèle, analyser le comportement de la forme fonctionnelle.

La comparaison des modèles BC_* et BC_A nous permet d'affirmer que $\hat{\lambda}_y$ (1.08) n'est pas statistiquement différent de un[3] et demeure par conséquent presque inchangé. Nous avons donc un paramètre généré qui ne constitue pas un

---

3.- Le test-t de student, quant à lui, nous permet de dire que $\hat{\lambda}_y = 1.08$ est statistiquement différent de 1.

TABLEAU 4

Valeurs des paramètres λ_y, λ_{xk}, ρ_ℓ, λ_{em}, δ_m, avec $X_m \equiv$ *RSIDCB* et valeurs correspondantes des fonctions de vraisemblance. - Séries * et A -.

Paramètres	Modèles				
	BC_*	BC_*HG_1	BC_A	BC_AHG_1	BC_AHC_1
λ_y	1.00	1.00	1.08 (55.43) (4.11)	1.49 (11.86) (3.93)	1.52 (82.67) (28.21)
λ_x [a]	1.00	1.00	1.08	1.49	1.52
ρ_1	0.00	0.00	0.00	0.00	0.00
ρ_2	0.00	0.00	0.00	0.00	0.00
ρ_{12}	0.00	0.00	0.00	0.00	0.00
λ_{e1}	0.00	-4.52 (-6.32)[b] (-7.72)[c]	0.00	-1.91 (- 9.80) (-14.94)	0.00
λ_{e2}	0.00	0.00	0.00	0.00	0.00
δ_1	0.00	288.15 (6.37)[b]	0.00	35.98 (13.28)	2.00
δ_2	0.00	0.00	0.00	0.00	0.00
$L(\omega^{**})$	147.65	149.38	147.99	155.60	152.15

a. Pour $\lambda_{x(1)}$, $\lambda_{x(2)}$, $\lambda_{x(3)}$, $\lambda_{x(4)}$, $\lambda_{x(5)}$.

b. La valeur du test statistique-t de student sous l'hypothèse que le paramètre est égal à zéro => $\hat{t}(0)$.

c. La valeur du test statistique-t de student sous l'hypothèse que le paramètre est égal à un => $\hat{t}(1) = \hat{t}(0) - \hat{t}(0)/\hat{\beta}$. Ce test est calculé pour λ_y, λ_x, λ_{em}.

UNIVERSITÉ DE MONTRÉAL

outil suffisant de correction. Une correction complète de l'hétéroscédasticité aurait exigé que $\hat{\lambda}_y$ prenne une valeur telle que la fonction de vraisemblance devienne approximativement équivalente à 149.38. Or, nous notons en appliquant le test T.3 que la valeur de la fonction de vraisemblance pour le modèle BC_A est plus petite que celle de BC_*HG_1 (147.99 < 149.38).

c) Si nous laissons estimer les paramètres de la partie fixe conjointement avec ceux de la partie stochastique, nous notons une augmentation de la valeur de la fonction de vraisemblance de 147.99 à 155.60, soit une nette amélioration. Cette augmentation de 7.61 est significative à un niveau de .0005 pour 1 degré de liberté. En sus, nous constatons une modification sensible de la forme fonctionnelle. D'une forme fonctionnelle linéaire, nous passons à une forme fonctionnelle qui élève les variables transformées à une puissance de un et demi. Cette augmentation dans la valeur du paramètre λ est responsable d'une partie de l'augmentation de $L(\omega^{**})$. Cependant, l'augmentation de $L(\omega^{**})$ est également dûe aux paramètres λ_{e1} et δ_1.

Nous avons ici un exemple probant que la forme fonctionnelle peut introduire de l'hétéroscédasticité en modifiant le rapport des variances.

Nous pouvons affirmer que le modèle est vraiment non-linéaire (en l'absence de l'autocorrélation des erreurs). De plus, il y a une forte présomption à croire que la forme fonctionnelle n'est pas robuste à l'hétéroscédasticité, car elle est modifiée selon qu'il y ait hétéroscédasticité ou non.

Une comparaison des modèles BC_AHG_1 et BC_AHC_1, nous permet de noter ici que la forme classique de l'hétéroscédasticité s'avère un outil de correction utile mais non suffisant. La valeur du test statistique pour la comparaison de ces deux modèles s'élève à 3.45; nous pouvons donc rejeter pour 2 degrés de liberté le modèle qui utilise la formulation classique de l'hétéros-

cédasticité avec confiance à un niveau de 95%. Cependant, l'utilisation de cette forme d'hétéroscédasticité nous aurait également permis de conclure à la non-linéarité du modèle. D'ailleurs la forme fonctionnelle entre ces deux modèles demeurant pratiquement inchangée, l'écart de 3.45 entre les valeurs des fonctions de vraisemblance vient confirmer une présence accrue de l'hétéroscédasticité dans le non-linéaire.

Nous allons maintenant analyser le comportement de la forme fonctionnelle lorsque nous libérons les contraintes sur les éléments vectoriels ρ_ℓ.

Cas 2

Les résultats générés pour chacun des modèles sont inscrits dans le tableau 5.

A l'intérieur des différentes applications, nous constatons que la principale contribution à l'augmentation de la valeur du maximum de vraisemblance est intrinsèquement reliée à la forte présence de l'autocorrélation des erreurs résiduelles. La valeur du test statistique pour la comparaison des modèles BC_* versus BC_*AU, s'élève à 24.87; ceci implique qu'avec les 3 degrés de liberté requis par les éléments non-nuls de ρ, l'hypothèse nulle que $\rho = 0$ est rejetée avec un haut niveau de confiance.

a) L'application du test T.10 pour les modèles BC_*AU et BC_*AUHG_1, nous permet de détecter, comme dans le cas 1, la présence de l'hétéroscédasticité; mais, tel que nous l'avons anticipé, du moins dans le linéaire, sa présence demeure marginale relativement à la présence de l'autocorrélation.

Néanmoins, nous pouvons considérer comme extrêmement intéressante, l'amélioration réalisée par la fonction de vraisemblance lorsque nous introduisons la forme générale de l'hétéroscédasticité. Les résultats passent de

TABLEAU 5

Valeurs des paramètres λ_y, λ_{xk}, ρ_ℓ, λ_{em}, δ_m, avec $X_m \equiv$ *RSIDCB* et valeurs correspondantes des fonctions de vraisemblance. - Séries ⋆ et A -.

Paramètres	Modèles					
	$BC_\star$	$BC_\star AU$	$BC_\star AUHG_1$	$BC_A AU$	$BC_A AUHG_1$	$BC_A AUHC_1$
λ_y	1.00	1.00	1.00	.62 (29.27) (-17.97)	.87 (5.04) (- .77)	1.27 (65.86) (13.81)
λ_x	1.00	1.00	1.00	.62	.87	1.27
ρ_1	0.00	.16 (2.48)	.18 (2.76)	.17 (2.58)	.18 (2.74)	.19 (2.95)
ρ_2	0.00	.22 (3.09)	.21 (3.07)	.19 (2.73)	.20 (2.97)	.20 (3.17)
ρ_{12}	0.00	.35 (6.43)	.38 (7.49)	.40 (8.10)	.39 (7.74)	.37 (6.97)
λ_{e1}	0.00	0.00	-.62 (-1.74) (-4.52)	0.00	-.83 (-1.85) (-4.07)	0.00
λ_{e2}	0.00	0.00	0.00	0.00	0.00	0.00
δ_1	0.00	0.00	2.70 (3.87)	0.00	2.88 (2.53)	2.00
δ_2	0.00	0.00	0.00	0.00	0.00	0.00
$L(\omega^{**})$	147.65	172.52	178.71	176.96	178.97	175.09

UNIVERSITÉ DE MONTRÉAL

172.52 à 178.71, soit une amélioration nettement significative de 6.19. La valeur des paramètres autorégressifs demeurant presque inchangée, nous pouvons conclure que cette augmentation est attribuée presque essentiellement à la présence de l'hétéroscédasticité.

b) Dans l'étape où nous généralisons la transformation Box-Cox sur les variables en laissant estimer les paramètres de la partie fixe tout en leur imposant une contrainte d'égalité, nous notons une amélioration sensible de la valeur du maximum de vraisemblance, celle-ci passant de 172.52 à 176.96, soit une différence de 4.44. De plus, la diminution de la valeur de $\hat{\lambda}_y$, qui passe de 1.00 à .62, témoigne de la capacité de la transformation Box-Cox à minimiser l'hétéroscédasticité. Cependant, la légère infériorité de la valeur de $L(\omega^{**})$ pour le modèle $BC_A AU$ comparativement à celle pour le modèle $BC_* AUHG_1$ (176.96 < 178.71) nous permet d'anticiper encore une présence réelle de l'hétéroscédasticité ou du moins nous permet de dire que la correction de l'hétéroscédasticité par $\hat{\lambda}_y$ n'est pas complète.

Nous concluons donc ici que la forme fonctionnelle est faussée puisqu'elle a joué un rôle de compensation pour l'hétéroscédasticité.

c) En introduisant la formulation générale de l'hétéroscédasticité, nous allons libérer le paramètre λ_y de son rôle de compensation pour l'hétéroscédasticité puisque nous ajoutons un instrument spécifique à cette fin. Les résultats induits de la détermination conjointe de la forme fonctionnelle des parties systématique et stochastique du modèle témoignent de l'apport véritable du modèle général. L'application du test du rapport des vraisemblances nous permet de rejeter à un niveau de confiance de 95% pour 1 degré de liberté, le modèle $BC_A AU$ (la valeur statistique pour T.13 étant égale à 2.01). Ceci signifie donc qu'il est nécessaire de corriger pour l'hétéroscédasticité avec la formulation générale. La valeur des paramètres λ_{em} et δ_m nous indique

également la nécessité de corriger pour l'hétéroscédasticité. Une correction avec la formulation classique implique, comme nous pouvons le constater dans le tableau 5, une diminution de la valeur de la fonction de vraisemblance de 3.88 par rapport à celle du modèle qui corrige avec la formulation générale de l'hétéroscédasticité. Cette différence nous permet donc de rejeter avec un degré élevé de confiance la formulation classique de l'hétéroscédasticité. D'ailleurs, une comparaison de BC_AAUHC_1 à BC_AAU révèle qu'il est même préférable de ne pas corriger du tout plutôt que d'utiliser la forme précédemment citée.

La comparaison des valeurs du maximum de vraisemblance des modèles BC_*AUHG_1 et BC_AAUHG_1, ne nous permet pas de rejeter le modèle BC_*AUHG_1, c'est-à-dire une forme fonctionnelle linéaire. Aussi, nous pouvons affirmer que dans le cas qui nous occupe, la présence de l'hétéroscédasticité et non pas la non-linéarité est la force qui agit sur la forme fonctionnelle en (b).

Il y a une présomption à croire que la forme fonctionnelle n'est pas robuste à l'hétéroscédasticité.

Nous allons maintenant analyser les implications sur la fonction de vraisemblance et sur la forme fonctionnelle si nous introduisons la variable *EIQP* comme un facteur susceptible d'introduire de l'hétéroscédasticité.

Cas 3

Les résultats générés pour chacun des modèles sont inscrits dans le tableau 6.

Nous limitons notre analyse aux modèles qui présentent l'adjonction des paramètres ρ_ℓ.

Nous remarquons tout d'abord que les résultats obtenus génèrent des

TABLEAU 6

Valeurs des paramètres λ_y, λ_{xk}, ρ_ℓ, λ_{em}, δ_m, avec $X_m \equiv EIQP$ et valeurs correspondantes des fonctions de vraisemblance. - Séries * et A -.

Paramètres	Modèles				
	BC_*AU	BC_*AUHG_2	BC_AAU	BC_AAUHG_2	BC_AAUHC_2
λ_y	1.00	1.00	.62 (29.27)	.89 (19.78)	.90 (47.48)
λ_x	1.00	1.00	(-17.97) .62	(-2.39) .89	(-5.10) .90
ρ_1	.16 (2.48)	.16 (2.52)	.17 (2.58)	.16 (2.51)	.18 (2.77)
ρ_2	.22 (3.09)	.22 (3.01)	.19 (2.73)	.21 (2.92)	.21 (2.88)
ρ_{12}	.35 (6.43)	.38 (7.78)	.40 (8.10)	.38 (7.91)	.37 (6.77)
λ_{e1}	0.00	0.00	0.00	0.00	0.00
λ_{e2}	0.00	11.69 (2.04)	0.00	13.02 (1.32)	0.00
δ_1	0.00	(1.86) 0.00	0.00	(1.22) 0.00	0.00
δ_2	0.00	.07 (.50)	0.00	.03 (.26)	2.00
$L(\omega^{**})$	172.52	178.01	176.96	178.13	171.53

UNIVERSITÉ DE MONTRÉAL

situations tout à fait identiques à celles obtenues dans le cas 2, pour les étapes (a) et (b). En effet, nous pouvons noter une présence marquée de l'hétéroscédasticité introduite par cette variable dans le linéaire (la valeur statistique pour le test T.10 est égale à 5.49, ce qui nous permet de rejeter avec un niveau de confiance de 99.9%, pour 1 degré de liberté, la présence d'erreurs homoscédastiques). De plus, nous pouvons noter que le paramètre $\hat{\lambda}_y$ (.62) réussit à compenser ici la presque totalité de l'hétéroscédasticité présente dans le modèle linéaire, et à changer par conséquent le caractère hétéroscédastique des résidus. La comparaison des valeurs du maximum de vraisemblance entre les modèles BC_AAU et BC_*AUHG_2 témoigne pour ce cas de l'apport considérable de la transformation Box-Cox sur la partie fixe comme outil de minimisation de l'hétéroscédasticité. La forme fonctionnelle s'avère donc faussée. De plus, la valeur statistique pour la comparaison des modèles BC_AAU et BC_AAUHG_2 s'élève à 1.17, ce qui nous permet de rejeter à un niveau de 90% le modèle BC_AAUHG_2. Par conséquent, il peut se faire un choix arbitraire entre le modèle qui utilise la procédure Box-Cox sur la partie systématique et le modèle qui associe la formulation générale de l'hétéroscédasticité à une forme fonctionnelle approximativement linéaire.

Les résultats inhérents au cas 3, nous permettent d'affirmer que le modèle ne semble pas robuste à la linéarité et de supposer que la forme fonctionnelle n'est pas robuste à l'hétéroscédasticité.

b) Deux variables prises conjointement.

Jusqu'ici, notre investigation s'est limitée à une prise en considération des variables *RSIDCB* et *EIQP*, comme sources possibles de l'hétéroscédasticité, lorsqu'elles sont considérées séparément. Nous allons à présent, analyser l'impact sur la fonction de vraisemblance et sur la forme fonctionnelle lorsque ces deux variables sont introduites conjointement dans $f(X_t)$.

Les résultats inscrits au tableau 7 laissent voir une augmentation extrêmement faible de la valeur du maximum de vraisemblance lorsque les deux variables sont combinées.

TABLEAU 7

Valeurs des paramètres λ_y, λ_{xk}, ρ_ℓ, λ_{em}, δ_m, avec $X_m \equiv$ *RSIDCB*, *EIQP* et valeurs correspondantes des fonctions de vraisemblance.
- Séries A et AA -.

Paramètres	Modèles	
	$BC_{AA}AUHG_{1,2}$	$BC_{AA}AUHC_{1,2}$
λ_y	.86 (4.87) (- .81)	1.50 (82.97) (27.74)
λ_x	.86	1.50
ρ_1	.18 (2.69)	.22 (3.15)
ρ_2	.20 (2.91)	.20 (2.81)
ρ_{12}	.39 (7.76)	.34 (5.78)
λ_{e1}	-.81 (-1.27) (-2.84)	0.00
λ_{e2}	-1.69 (- .05) (- .09)	0.00
δ_1	2.98 (2.31)	2.00
δ_2	-.26 (-.37)	2.00
$L(\omega^{**})$	179.06	163.71
RSIDCB → λ_y	.87	1.27
$L(\omega^{**})$	178.97	175.09
EIQP → λ_y	.89	.90
$L(\omega^{**})$	178.13	171.51

D'une valeur de 178.13 avec *EIQP* ou d'une valeur de 178.97 avec *RSIDCB*, nous passons à une valeur de 179.06, soient des augmentations respectives de seulement .93 et de .09; ceci implique que pour deux degrés de liberté, nous pouvons rejeter l'hypothèse d'une plus grande présence de l'hétéroscédasticité lorsque ces deux variables sont introduites simultanément dans $f(X_t)$, et ce même à un niveau très marginal.

De plus, le tableau 7 indique que l'utilisation de la forme multiplicative multivariée classique de l'hétéroscédasticité entraîne une diminution notable de la valeur logarithmique de la fonction du maximum de vraisemblance.

4.2.2. Utilité de la généralité de la paramétrisation, à partir d'une analyse des différents modèles sur la base des hypothèses émises sur la partie fixe du modèle.

A ce point de notre étude, nous pouvons effectuer une comparaison des modèles sous l'hypothèse d'une plus grande généralité des paramètres de la partie fixe. L'importance à accorder à la relaxation des contraintes sur la partie fixe sera analysée à partir de la modification significative ou non significative de $L(\omega^{**})$ et à partir de la modification des paramètres λ_y et λ_{xk}. Les modèles BC_SAU et BC_SAUHG_1, pour S = *, A, B, C, D, E, nous instruiront à cet effet.

Les résultats pour chacun des deux modèles sont respectivement présentés dans les tableaux 8 et 9.

Les résultats inscrits dans les tableaux 10 et 11, indiquent, quant à eux, les valeurs des tests du rapport du maximum de vraisemblance pour chacune des paires d'hypothèses sur λ_y, $\lambda_{x(1)}$, $\lambda_{x(2)}$, $\lambda_{x(3)}$, $\lambda_{x(4)}$, $\lambda_{x(5)}$.

A l'intérieur des diverses applications de chacun des deux modèles, nous remarquons une augmentation des valeurs du maximum de vraisemblance, au

TABLEAU 8

Valeurs des paramètres λ_y, λ_{xk}, ρ_ℓ, λ_{em}, δ_m et valeurs correspondantes des fonctions de vraisemblance. Modèle $BC_S AU$ - Séries *, A, B, C, D, E. -

Paramètres	Modèles						
	BC_*AU	BC_*AU	BC_AAU	BC_BAU	BC_CAU	BC_DAU	BC_EAU
λ_y	1.00	0.00	.62 (29.27) (-17.97)	.60 (28.37) (-18.62)	.61 (29.04) (-18.18)	.61 (28.34) (-18.29)	.66 (29.13) (-15.25)
$\lambda_{x(1)}$	1.00	0.00	.62	.12 (.23) (- 1.65)	-1.90 (- 2.22) (- 3.38)	-2.12 (- 2.77) (- 4.08)	.24 (.20) (- .62)
$\lambda_{x(2)}$	1.00	0.00	.62	.12	-1.90	-2.12	1.48 (.53) (.17)
$\lambda_{x(3)}$	1.00	0.00	.62	.12	-1.90	-2.12	-2.58 (- 2.31) (- 3.21)
$\lambda_{x(4)}$	1.00	0.00	.62	.12	.34 (.64) (- 1.22)	.31 (.58) (- 1.30)	.47 (1.00) (- 1.11)
$\lambda_{x(5)}$	1.00	0.00	.62	.12	-1.90	8.14 (1.04) (.91)	6.53 (.97) (.82)
ρ_1	.16 (2.48)	.18 (3.10)	.17 (2.58)	.17 (2.61)	.15 (2.25)	.14 (2.03)	.17 (2.45)
ρ_2	.22 (3.09)	.08 (1.59)	.19 (2.73)	.18 (2.67)	.18 (2.57)	.17 (2.55)	.20 (2.92)
ρ_{12}	.35 (6.43)	.72 (18.74)	.40 (8.10)	.42 (8.12)	.40 (8.07)	.41 (8.37)	.33 (6.41)
λ_{e1}	0.00	0.00	0.00	0.00	0.00	0.00	0.00
λ_{e2}	0.00	0.00	0.00	0.00	0.00	0.00	0.00
δ_1	0.00	0.00	0.00	0.00	0.00	0.00	0.00
δ_2	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$L(\omega^{**})$	172.52	175.11	176.96	177.51	178.59	179.14	181.69
Nombre de paramètres	3	3	4	5	6	7	9

TABLEAU 9

Valeurs des paramètres λ_y, λ_{xk}, ρ_ℓ, λ_{em}, δ_m et valeurs correspondantes des fonctions de vraisemblance. Modèle $BC_S AUHG_1$ - Séries *, A, B, C, D, E. -

Paramètres	Modèles						
	$BC_* AUHG_1$	$BC_* AUHG_1$	$BC_A AUHG_1$	$BC_B AUHG_1$	$BC_C AUHG_1$	$BC_D AUHG_1$	$BC_E AUHG_1$
λ_y	1.00	0.00	.87 (5.04)	.84 (4.79)	.78 (5.32)	.76 (4.96)	.77 (5.56)
$\lambda_{x(1)}$	1.00	0.00	.87 (- .77)	(- .93)	(-1.48)	(-1.53)	(- 1.69)
$\lambda_{x(2)}$	1.00	0.00	.87	.64 (1.23)	-1.83 (-2.25)	-1.99 (-2.59)	.53 (.45)
$\lambda_{x(3)}$	1.00	0.00	.87	.64	(-3.47) -1.83	(-3.89) -1.99	(- .39) 1.55
$\lambda_{x(4)}$	1.00	0.00	.87	.64	-1.83	-1.99	(.55) (.19) -2.33
$\lambda_{x(5)}$	1.00	0.00	.87	.64	-1.83	.79 (1.15)	(-1.92) (- 2.75) .72 (1.08)
ρ_1	.18 (2.76)	.23 (3.72)	.18 (2.74)	.18 (2.77)	.17 (2.40)	(- .31) .15 (2.20)	(- .43) .18 (2.62)
ρ_2	.21 (3.07)	.03 (.57)	.20 (2.97)	.20 (2.88)	.19 (2.71)	.18 (2.80)	.21 (3.10)
ρ_{12}	.38 (7.49)	.68 (18.59)	.39 (7.74)	.40 (7.58)	.39 (7.76)	.40 (7.90)	.33 (6.40)
λ_{e1}	-.62 (-1.74)	1.77 (2.85)	-.83 (-1.85)	-.87 (-1.75)	.03 (.03)	-.16 (- .19)	-.20 (- .27)
λ_{e2}	0.00	(-4.52) 0.00	(-4.07) 0.00	(-3.76) 0.00	(-1.16) 0.00	(-1.36) 0.00	(- 1.32) 0.00
δ_1	2.70 (3.87)	-.04 (-1.09)	2.88 (2.53)	2.77 (.24)	.52 (.76)	.65 (.81)	.64 (.78)
δ_2	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$L(\omega^{**})$	178.71	178.92	178.97	179.07	180.12	180.37	182.97
Nombre de paramètres	5	5	6	7	8	9	11

TABLEAU 10

Valeurs des tests statistiques du rapport du maximum de vraisemblance pour les paramètres λ_y, λ_{xk}. Modèle $BC_S AU$ - Séries *, A, B, C, D, E. -

$BC_* AU$[1]	$BC_* AU$[2]	$BC_A AU$	$BC_B AU$	$BC_C AU$	$BC_D AU$	$BC_E AU$
0.	2.59 (0 D.L.)	4.44[3] (1 D.L.)	- -	- -	- -	- -
		0.	.55 (1 D.L.)	1.63 (2 D.L.)	2.18 (3 D.L.)	4.73[4] (5 D.L.)
			0.	1.08 (1 D.L.)	1.63 (2 D.L.)	4.18[4] (4 D.L.)
				0.	.55 (1 D.L.)	3.10 (3 D.L.)
					0.	2.55[4] (2 D.L.)
						0.

1. Modèle linéaire.
2. Modèle logarithmique.
3. Statistiquement significatif à un niveau de confiance de 99.5%.
4. Statistiquement significatif à un niveau de confiance de 90.0%

TABLEAU 11

Valeurs des tests statistiques du rapport du maximum de vraisemblance pour les paramètres λ_y, λ_{xk}. Modèle $BC_S AUHG_1$ - Séries *, A, B, C, D, E. -

$BC_* AUHG_1$[1]	$BC_* AUHG_1$[2]	$BC_A AUHG_1$	$BC_B AUHG_1$	$BC_C AUHG_1$	$BC_D AUHG_1$	$BC_E AUHG_1$
0.	.21 (0 D.L.)	.26 (1 D.L.)	- -	- -	- -	- -
		0.	.10 (1 D.L.)	1.15 (2 D.L.)	1.40 (3 D.L.)	4.00 (4 D.L.)
			0.	1.05 (1 D.L.)	1.30 (2 D.L.)	3.90[4] (4 D.L.)
				0.	.25 (1 D.L.)	2.85 (3 D.L.)
					0.	2.60[4] (2 D.L.)
						0.

fur et à mesure que nous libérons les contraintes d'égalité sur les paramètres de la partie fixe.

Le modèle contraint linéaire ($\lambda_y = \lambda_{x(1)} = \lambda_{x(2)} = \lambda_{x(3)} = \lambda_{x(4)} = \lambda_{x(5)} = 1$), donne la plus faible valeur du maximum de vraisemblance. Une amélioration de 2.59 est réalisée pour le modèle $BC_S AU$, lorsque nous imposons la forme logarithmique ($\lambda_y = \lambda_{x(1)} = \lambda_{x(2)} = \lambda_{x(3)} = \lambda_{x(4)} = \lambda_{x(5)} = 0$), la valeur du maximum de vraisemblance passant de 172.52 à 175.11.[4] Pour cette étape, le modèle $BC_S AUHG_1$ ne présente qu'une augmentation de .21. Si nous maintenons l'égalité entre les différents paramètres tout en les laissant estimer ($\hat{\lambda}_y = \hat{\lambda}_{x(1)} = \hat{\lambda}_{x(2)} = \hat{\lambda}_{x(3)} = \hat{\lambda}_{x(4)} = \hat{\lambda}_{x(5)}$), nous notons pour le premier modèle, une augmentation de $L(\omega^{**})$ évaluée à 4.44, ce qui est statistiquement significatif pour 1 degré de liberté à un niveau de confiance de 99.5%; dans le second modèle, nous remarquons un accroissement de seulement .26.

Dans l'étape où la forme la plus générale est générée ($\hat{\lambda}_y \neq \hat{\lambda}_{x(1)} \neq \hat{\lambda}_{x(2)} \neq \hat{\lambda}_{x(3)} \neq \hat{\lambda}_{x(4)} \neq \hat{\lambda}_{x(5)}$), nous avons par opposition à celle où les lambdas sont contraints à varier conjointement, une amélioration de 4.73 (pour $BC_S AU$), ce qui pour 5 degrés de liberté est statistiquement significatif à un niveau de confiance de .10. Pour le modèle $BC_S AUHG_1$, l'amélioration n'est pas significative, même à ce niveau. Par ailleurs, nous notons avec intérêt que la relaxation de la contrainte d'égalité entre le paramètre λ_y et les paramètres λ_{xk} entraîne un accroissement minime de la valeur de $L(\omega^{**})$.

Un relâchement graduel des contraintes d'égalité entre les lambdas

---

4.- Dans le modèle où les éléments du vecteur ρ sont contraints à être égaux à zéro, soit BC_S pour S = *, A, B, C, D, E, nous notons une amélioration très nette de la valeur du maximum de vraisemblance, lorsque nous imposons une forme linéaire plutôt qu'une forme logarithmique, la valeur passant de 93.37 à 147.65, soit une différence accrue de 54.28. Une valeur plus élevée de $L(\omega^{**})$ pour le modèle logarithmique peut peut-être s'expliquer par un paramètre d'autocorrélation ρ_{12} beaucoup plus élevé .

laisse voir un accroissement non proportionnel relativement au nombre de degrés de liberté. (Les résultats sont montrés sur la diagonale principale dans chacun des deux derniers tableaux).

Les résultats obtenus pour ces deux modèles donnent un aperçu de la situation qui prévaut pour chacun des autres modèles non présentés. (Les améliorations de la valeur du maximum de vraisemblance entre les différentes séries s'avèrent pour certains modèles encore plus marginales.)

Une analyse des paramètres $\hat{\lambda}$ à travers chacune des applications pour un même modèle, nous permet de noter la stabilité relative de la forme fonctionnelle du côté gauche de l'équation. Il est intéressant de noter que les valeurs du test statistique du rapport du maximum de vraisemblance entre les modèles pris deux à deux ne sont pas statistiquement significatives, même si les paramètres qui modifient les variables originales $y_{.,}$ varient beaucoup. Aussi, le paramètre λ_y influence beaucoup plus la valeur du maximum de vraisemblance que ne le font les paramètres λ_{xk}. Ceci est sans doute expliqué par la présence du paramètre λ_y dans le Jacobien de la transformation de $y_t^{(\lambda_y)}$ au y_t observé.

4.2.3. Détermination du modèle qui maximise $L(\omega^{**})$.

Nous allons mettre un point à notre discussion de l'estimation de la forme fonctionnelle en présentant les résultats dérivés du modèle le plus général, $BC_{EE}AUHG_{1,2}$.

Les paramètres estimés pour le modèle qui maximise $L(\omega^{**})$ sont présentés dans le tableau 12.

Le modèle le plus général (13 paramètres) nous permet de conclure que la forme fonctionnelle qui donne le "meilleur" ajustement pour la fonction de

TABLEAU 12

Valeurs des paramètres λ_y, λ_{xk}, ρ_ℓ, λ_{em}, δ_m du modèle qui optimise $L(\omega^{**})$.

Paramètres	Modèle $BC_{EE}^{AUHG}{}_{1,2}$
λ_y	.84 (5.07) (- .94)
$\lambda_{x(1)}$	.50 (.39) (- .39)
$\lambda_{x(2)}$	1.65 (.57) (.22)
$\lambda_{x(3)}$	-2.30 (-1.74) (-2.49)
$\lambda_{x(4)}$	.87 (1.46) (- .22)
$\lambda_{x(5)}$	5.46 (.80) (.65)
ρ_1	.18 (2.58)
ρ_2	.21 (2.97)
ρ_{12}	.33 (6.36)
λ_{e1}	-.59 (- .34) (- .91)
λ_{e2}	15.97 (.48) (.45)
δ_1	.85 (.60)
δ_2	.005 (.07)
$L(\omega^{**})$	183.22

demande d'essence dans la province de Québec n'est pas doublement linéaire. En effet, même si $\hat{\lambda}_y$ est approximativement linéaire, les lambdas sur les variables explicatives sont tous (sauf $\hat{\lambda}_{x(4)}$) différents du linéaire, même si les tests $\hat{t}(1)$ qui leur sont associés ne le justifient pas.

Cependant, la valeur du test chi-carré (4.16 < 4.62) pour un niveau de confiance de 90% avec 5 degrés de liberté ne justifie pas le rejet du modèle $BC_{AA}AUHG_{1,2}$, la valeur de $L(\omega^{**})$ s'élevant,tel que montré dans le tableau 7, à 179.06 pour ce dernier modèle. Nous pouvons donc accepter le modèle $BC_{AA}AUHG_{1,2}$ dans lequel la forme fonctionnelle est approximativement doublement linéaire ($\hat{\lambda}_y = \hat{\lambda}_{xk} = .86$).

4.3. Effets de la forme fonctionnelle et de la spécification de l'erreur.

L'estimation de la forme fonctionnelle de la partie fixe du modèle à l'aide de la transformation Box-Cox peut induire des modifications sensibles sur la grandeur et la signification des coefficients β_k et par conséquent, sur les élasticités dérivées de la demande et sur les tests statistiques-t de student. Une mauvaise spécification de la forme fonctionnelle de la partie fixe peut induire des résultats qui sont moins raisonnables et moins significatifs que ceux obtenus des modèles qui incorporent une formulation souple de la forme fonctionnelle. De même, la transformation Box-Cox appliquée à la partie stochastique peut engendrer des variantes dans les résultats induits, les valeurs des paramètres λ de la partie fixe n'étant pas indépendantes des valeurs des paramètres qui spécifient la distribution des résidus,tel que déjà noté.

4.3.1. Effets de la forme fonctionnelle et de la spécification de l'erreur sur la distribution des résidus autorégressifs.

Nous avons déjà analysé d'une manière exhaustive l'effet de la forme fonctionnelle sur la distribution des résidus hétéroscédastiques (introduction, correction de l'hétéroscédasticité). Nous allons maintenant, à l'aide de quatre de nos modèles, étudier l'impact de la forme fonctionnelle de la partie fixe et/ou de la partie stochastique du modèle sur les résidus autorégressifs.

La différence des valeurs logarithmiques des fonctions de vraisem-

blance entre les modèles BC_* ($\lambda_y = \lambda_{xk} = 1$) et BC_A s'élève à .34 ($\hat{L} = -147.65 + 147.99$) ce qui indique que, pour 1 degré de liberté, les ordres du schéma autorégressif postulés à $\lambda_y = \lambda_{xk} = 1$; $\rho = 0$; $\lambda_e = 0$; $\delta = 0$ ne sont pas statistiquement significatifs de ceux qui auraient pu être postulés à partir de $\hat{\lambda}$; $\rho = 0$; $\lambda_e = 0$; $\delta = 0$. Une analyse des fonctions d'autocorrélation et d'autocorrélation partielle présentées dans le tableau 13, confirme l'affirmation précédente. En effet, si nous passons du modèle linéaire en A à une forme fonctionnelle approximativement linéaire en C, ($\hat{\lambda}_y = \hat{\lambda}_{xk} = 1.08$), il n'y a pas de changement significatif dans la structure. Toutefois, si nous prenons en considération la corrélation temporelle des erreurs, en B, et que nous postulons une spécification complexe de la structure autorégressive ($\rho = \rho_1, \rho_2, \rho_{12}$), de sorte que les délais les plus faibles sont affectés par ces paramètres, nous constatons que l'autocorrélation est corrigée. Ceci signifie que les tests statistiques-t ne sont pas biaisés vers le haut. Une comparaison des valeurs des fonctions entre les modèles BC_* et BC_*AU nous permet, tel que noté antérieurement, de rejeter avec un très haut niveau de confiance pour 3 degrés de liberté, l'hypothèse que $\rho = 0$. Aussi, un changement notable de la structure était prévu. Dans le modèle le plus général et qui optimise $L(\omega^{**})$, en D, une augmentation de 10.7 (significative à un niveau de confiance de 95% pour 10 degrés de liberté) de la valeur de la fonction de vraisemblance par rapport au modèle BC_*AU, en B, nous permet également d'anticiper une modification de la structure autorégressive. Nous constatons avec intérêt que celle-ci n'est pas modifiée radicalement même si nous avons introduit une forme d'hétéroscédasticité et même si la forme fonctionnelle du côté droit de l'équation est sensiblement modifiée. Dans aucun de nos modèles, nous avons trouvé des changements drastiques de la structure sauf dans le modèle logarithmique BC_*. Nonobstant, elle nous aurait quand même permis de postuler le même processus autorégressif. Nous pouvons ajouter que dans BC_*AU et BC_*AUHG_1 (log), l'autocorrélation des 2° et 12° ordres n'est pas complètement corrigée.

TABLEAU 13

Fonctions d'autocorrélation et d'autocorrélation partielle, en différence des résidus d'ordre 0.

A. $\lambda_y=\lambda_{x(1)}=\lambda_{x(2)}=\lambda_{x(3)}=\lambda_{x(4)}=\lambda_{x(5)}=1$; $\rho=\rho_1=\rho_2=\rho_{12}=0$; $\lambda_e=0$; $\delta=0$.

Délais													Ecart-type estimé pour la rangée
						Autocorrélations							
1-12	.21	.18	.06	.04	.06	.09	.03	.04	-.01	-.09	.02	.23	(.07)
13-24	-.06	-.07	-.06	-.09	-.07	-.02	-.08	.01	-.07	-.05	-.07	.13	(.07)
25-36	-.02	.06	.01	.02	-.04	-.07	-.01	-.04	-.05	-.06	-.03	.15	(.08)
37-48	.06	-.07	.06	-.06	.02	-.08	-.01	-.05	-.07	-.06	-.09	.16	(.08)
49-60	-.02	-.06	.00	-.06	-.08	-.06	-.05	-.00	-.05	-.07	-.08	.04	(.08)
						Autocorrélations partielles							
1-12	.21	.14	-.00	.01	.05	.07	-.02	.01	-.03	-.10	.06	.26	(.07)

B. $\lambda_y=\lambda_{x(1)}=\lambda_{x(2)}=\lambda_{x(3)}=\lambda_{x(4)}=\lambda_{x(5)}=1$; $\hat{\rho}=\hat{\rho}_1, \hat{\rho}_2, \hat{\rho}_{12}$; $\lambda_e=0$; $\delta=0$.

						Autocorrélations							
1-12	.05	.08	.00	-.07	-.00	.03	-.03	-.00	.02	-.10	.13	-.03	(.07)
13-24	-.11	-.11	.02	-.09	-.05	-.02	-.11	.01	-.04	.05	-.04	.04	(.07)
25-36	.01	.15	-.02	.03	-.09	-.09	-.04	-.06	.01	-.02	.02	.08	(.07)
37-48	.11	-.06	.09	-.09	.06	-.08	-.00	-.03	-.02	.04	-.04	.17	(.07)
49-60	-.03	.03	.02	-.00	-.10	-.02	-.05	.03	.01	-.01	.01	-.00	(.08)
						Autocorrélations partielles							
1-12	.05	.07	-.01	-.07	.00	.04	-.03	-.01	.03	-.10	.13	-.03	(.07)

UNIVERSITÉ DE MONTRÉAL

C. $\hat{\lambda}_y=\hat{\lambda}_{x(1)}=\hat{\lambda}_{x(2)}=\hat{\lambda}_{x(3)}=\hat{\lambda}_{x(4)}=\hat{\lambda}_{x(5)}=1.08$; $\rho=\rho_1=\rho_2=\rho_{12}=0$; $\lambda_e=0$; $\delta=0$.

Délais					Autocorrélations								Ecart-type estimé pour la rangée
1-12	.20	.17	.05	.05	.07	.09	.03	.04	-.01	-.10	.02	.22	(.07)
13-24	-.07	-.08	-.07	-.09	-.06	-.02	-.07	.02	-.07	-.05	-.07	.12	(.07)
25-36	-.02	.06	.01	.02	-.03	-.07	-.01	-.03	-.05	-.05	-.03	.15	(.08)
37-48	.06	-.07	.06	-.07	.02	-.09	-.01	-.04	-.07	-.06	-.09	.15	(.08)
49-60	-.02	-.06	.00	-.06	-.08	-.06	-.04	.00	-.04	-.07	-.08	.04	(.08)
					Autocorrélations partielles								
1-12	.20	.13	-.00	.01	.05	-.07	-.02	.01	-.03	-.11	.06	.25	(.07)

D. $\hat{\lambda}_y\neq\hat{\lambda}_{x(1)}\neq\hat{\lambda}_{x(2)}\neq\hat{\lambda}_{x(3)}\neq\hat{\lambda}_{x(4)}\neq\hat{\lambda}_{x(5)}$; $\hat{\rho}=\hat{\rho}_1, \hat{\rho}_2, \hat{\rho}_{12}$; $\hat{\lambda}_{e1}, \hat{\lambda}_{e2}$; $\hat{\delta}_1, \hat{\delta}_2$.

					Autocorrélations								
1-12	.05	.07	-.01	-.04	.02	-.00	-.00	-.02	.01	-.08	.06	-.00	(.07)
13-24	-.10	-.10	.04	-.09	-.06	-.05	-.07	.00	-.02	.04	-.08	.05	(.07)
25-36	-.05	.10	-.00	.01	-.09	-.14	-.05	-.07	-.01	-.04	-.01	.11	(.07)
37-48	.11	-.04	.11	-.09	.12	-.09	.01	-.04	-.01	.06	-.02	.19	(.07)
49-60	-.04	.03	.01	-.02	-.03	-.01	-.02	.07	.01	.01	.01	-.00	(.08)
					Autocorrélations partielles								
1-12	.05	.07	-.02	-.04	.03	.00	-.01	-.02	.02	-.08	.06	.00	(.07)

UNIVERSITE DE MONTREAL

. Coefficient de corrélation multiple.

Le tableau 14 fournit trois mesures de coefficient de corrélation multiple. Ces trois mesures sont présentées dans les expressions (3.1), (3.3) et (3.4) de l'appendice 3.

De par sa définition, le premier R^2 nous donne une indication du "goodness of fit". Chacun d'eux étant supérieur à .950, nous avons donc une très bonne qualité de l'ajustement du modèle à la réalité statistique. Pour chacun des modèles présentés, la première mesure de R^2 est supérieure à celle des deux autres. (Dans les cas non montrés, une seule leur est inférieure). En ce qui concerne les deux autres mesures, elles sont moins convaincantes de la mesure de l'erreur du modèle puisqu'elles sont calculées en termes de la valeur transformée de la variable dépendante, tel qu'exprimé en annexe. Elles sont toutes assez élevées, sauf pour le modèle BC_*Au lorsque $\lambda_y = \lambda_{xk} = 0$; dans ce dernier cas, $R^2 = .195$.

Il est par conséquent surprenant de constater que $R^2(E(y))$ demeure pratiquement inchangé même si moins de contraintes sont imposées sur les paramètres des parties fixe et stochastique. Ainsi, la plus grande variation dans le coefficient de détermination survient si nous passons du modèle BC_AHG_1 à BC_AAUHG_1, celui-ci passant de .976 à .980. (Nous ne considérons pas ici les modèles linéaire et logarithmique). De plus, il est surprenant de constater que même si $R^2(E(y))$ est légèrement inférieur dans le modèle BC_AHG_1, la variance résiduelle est quand même inférieure celle-ci étant de 4.54E-8 contre .010. De plus, même si la valeur du coefficient de détermination est la même entre deux modèles particuliers, il peut y avoir quand même une variation sensible entre leur variance résiduelle.

4.3.2. Effets de la forme fonctionnelle et de la spécification de l'erreur sur la signification des variables.

TABLEAU 14

Valeurs des coefficients paramétriques, des élasticités, des tests statistiques et de $L(\omega^{**})$ pour divers modèles.

		$BC_\bullet$		$BC_\bullet AU$		$BC_\bullet$		$BC_\bullet AU$	
Code des paramètres		c	t	c	t	c	t	c	t
λ_y		0.00		0.00		1.00		1.00	
$\lambda_{x(1)}$		0.00		0.00		1.00		1.00	
$\lambda_{x(2)}$		0.00		0.00		1.00		1.00	
$\lambda_{x(3)}$		0.00		0.00		1.00		1.00	
$\lambda_{x(4)}$		0.00		0.00		1.00		1.00	
$\lambda_{x(5)}$		0.00		0.00		1.00		1.00	
ρ_1		0.00		.18	3.10	0.00		.16	2.48
ρ_2		0.00		.08	1.59	0.00		.22	3.09
ρ_{12}		0.00		.72	18.74	0.00		.35	6.43
λ_{e1}		0.00		0.00		0.00		0.00	
λ_{e2}		0.00		0.00		0.00		0.00	
δ_1		0.00		0.00		0.00		0.00	
δ_2		0.00		0.00		0.00		0.00	
Code	Nom des variables	n	t	n	t	n	t	n	t
FADCB	Tarif réel du t.e.c.	.18	2.70	.01	.17	.11	2.27	.08	1.05
PGRMD	Coût réel de l'essence	.01	.04	-.16	-1.51	.05	.50	-.05	-.51
NCQR	Nombre d'automobiles	.50	3.36	-.29	-1.11	.53	4.28	.54	3.03
PCB	Indice des prix à la cons.	-.06	-.38	.09	.25	-.20	-2.02	-.20	-1.18
RAWEQ	Salaire réel moyen	.52	1.84	-.19	-.43	.64	3.22	.84	2.84
TW	Temps d'attente du t.e.c.	.40	2.93	.04	.25	.30	2.90	.49	3.25
TT	Temps de transport en t.e.c.	.07	.71	.08	1.15	.05	.68	-.00	-.03
TAUTO	Temps de transport en auto	-.34	-1.89	.12	.31	-.12	-.88	-.70	-1.52
TLK	Inc. dû à la température	-.00	-.68	.00	.99	-.02	-1.85	-.01	-.72
RFK	Inc. dû à la pluie	.00	.36	-.00	-1.12	-.00	-.52	-.01	-1.21
SFK	Inc. dû à la neige	.00	.71	.00	.76	.01	1.07	.00	.55
CSFMR	Inc. dû à la neige acc.	-.00	-6.26	-.00	-1.55	-.03	-5.49	-.02	-3.61
RCUTPN	Ind. prés. scol. (pr. & sec.)	.00	.73	-.00	-1.32	-.01	-.72	-.02	-1.98
ISECON	Ind. prés. scol. (coll. & un.)	-.00	-1.16	.00	.02	-.05	-2.83	-.02	-.83
RSIDCB	Ind. de déplac.-magasinage	.10	3.69	.07	1.93	.11	6.00	.10	5.36
EIQP	Niveau de l'emploi	-.00	-.05	.00	.03	.09	1.63	.03	.33
EX7	Expo'67	.00	1.84	.00	.03	.00	2.33	.00	1.19
STR	Grève à la C.T.C.U.M.	.00	.97	.00	1.90	.00	1.95	.00	1.90
VWM	Congé des travailleurs	.00	1.97	.00	.69	.01	2.96	.01	1.40
WD	Jours de travail/mois	.56	3.69	.30	1.12	.48	4.22	.45	2.74
SAT	Samedis/mois	.00	.08	-.01	-.14	-.01	-.19	-.03	-.60
SHD	Dim. & jrs. de fêtes/mois	.02	.40	.00	.58	-.00	-.01	-.00	-.04
C	Constante	.25	.35	4.86	.82	.49	1.22	-.15	-.30
$R^2(E(y))$		.966		.978		.976		.980	
$R^2(E(y^{**}))$		.962		.195		.976		.817	
$R^2(\hat{y}^{**})$		.962		.195		.976		.817	
Variance résiduelle		.005		.003		.285		.230	
$L(\omega^{**})$		93.37		175.11		147.65		172.52	

TABLEAU 14 (suite)

Valeurs des coefficients paramétriques, des élasticités, des tests statistiques et de $L(\omega^{**})$ pour divers modèles.

		BC_*AUHG_1		BC_AAU		BC_AHG_1		BC_AAUHC_1		BC_AAUHC_1	
	Code des paramètres	C	t	C	t	C	t	C	t	C	t
	λ_y	1.00		.62	29.27 -17.97	1.49	11.86 3.93	.87	5.04 -.77	1.27	65.86 13.81
	$\lambda_{x(1)}$	1.00		.62		1.49		.87		1.27	
	$\lambda_{x(2)}$	1.00		.62		1.49		.87		1.27	
	$\lambda_{x(3)}$	1.00		.62		1.49		.87		1.27	
	$\lambda_{x(4)}$	1.00		.62		1.49		.87		1.27	
	$\lambda_{x(5)}$	1.00		.62		1.49		.87		1.27	
	ρ_1	.18	2.76	.17	2.58	0.00		.18	2.74	.19	2.95
	ρ_2	.21	3.07	.19	2.73	0.00		.20	2.97	.20	3.17
	ρ_{12}	.38	7.49	.40	8.10	0.00		.39	7.74	.37	6.97
	λ_{e1}	-.62	-1.74 -4.52	0.00		-1.91	-9.80 -14.94	-.83	-1.85 -4.07	0.00	
	λ_{e2}	0.00		0.00		0.00		0.00		0.00	
	δ_1	2.70	3.87	0.00		35.98	13.28	2.88	2.53	2.00	
	δ_2	0.00		0.00		0.00		0.00		0.00	
Code	Nom des variables	n	t	n	t	n	t	n	t	n	t
FADCB	Tarif réel du t.e.c.	.11	1.55	.10	1.43	.13	1.63	.11	1.17	.14	2.06
PGRMD	Coût réel de l'essence	-.04	-.40	-.06	-.57	.14	1.28	-.05	-.45	-.01	-.11
NCQR	Nombre d'automobiles	.58	3.19	.59	2.97	.58	3.23	.59	2.42	.57	3.54
PCB	Ind. des prix à la cons.	-.15	-.76	-.20	-1.02	-.25	-2.75	-.18	-.85	-.06	-.32
RAWEQ	Salaire réel moyen	.76	2.31	.82	2.68	.52	2.48	.80	2.36	.58	1.81
TW	Temps d'attente du t.e.c.	.47	2.93	.53	3.25	.22	1.77	.50	2.94	.40	2.48
TT	Temps de trans. en t.e.c.	.01	.14	.01	.08	.04	.38	.01	.11	.03	.37
TAUTO	Temps de trans. en auto	-.70	-1.66	-.83	-1.81	-.06	-.36	-.78	-1.62	-.51	-1.43
TLK	Inc. dû à la température	-.01	-.99	-.00	-.48	-.08	-2.25	-.01	-.74	-.03	-1.48
RFK	Inc. dû à la pluie	-.01	-1.29	-.00	-1.13	-.03	-.99	-.01	-.99	-.02	-1.46
SFK	Inc. dû à la neige	.00	.76	.00	.64	.01	.97	.00	.64	.00	.82
CSFMR	Inc. dû à la neige acc.	-.02	-3.45	-.01	-3.59	-.06	-3.51	-.02	-2.15	-.03	-3.14
RCUTPN	Ind. pr. sc. (pr. & sec.)	-.03	-1.87	-.01	-1.64	-.05	-1.15	-.02	-1.36	-.07	-2.06
ISECON	Ind. pr. sc. (col. & un.)	-.02	-.82	-.01	-.53	-.19	-2.09	-.01	-.67	-.06	-1.23
RSIDCB	Ind. de cépla.-magasinage	.10	4.08	.11	4.62	.12	5.16	.10	3.86	.10	2.91
EIQP	Niveau de l'emploi	.03	.39	.01	.10	.23	2.13	.02	.24	.09	1.02
EX7	Expo'67	.00	.94	.00	1.14	.00	.56	.00	.92	.00	.72
STR	Grève à la C.T.C.U.M.	.00	1.71	.00	1.90	.00	.87	.00	1.48	.00	1.55
VWM	Congé des travailleurs	.01	1.61	.00	1.15	.05	2.57	.01	1.17	.02	2.22
WD	Jours de travail/mois	.29	1.95	.43	2.46	.17	1.97	.32	1.79	.18	1.64
SAT	Samedis/mois	-.04	-.95	-.03	-.58	-.03	-.78	-.04	-.91	-.04	-1.10
SHD	Dim. & jrs. de fêtes/mois	-.02	-.48	.00	.01	-.05	-2.02	-.02	-.37	-.03	-.83
C	Constante	.17	.35	-.24	-.52	1.57	2.03	.03	.07	.75	1.38
$R^2(E(y))$		.980		.980		.976		.980		.979	
$R^2(E(y^{**}))$		.795		.778		.952		.765		.895	
$R^2(\hat{y}^{**})$		.795		.778		.952		.765		.895	
Variance résiduelle		.014		.041		4.54E-8		.010		.028	
$L(\omega^{**})$		178.71		176.96		155.60		178.97		175.09	

TABLEAU 14 (suite)

Valeurs des coefficients paramétriques, des élasticités, des tests statistiques et de $L(\omega^{**})$ pour divers modèles.

		BC_AAUHG_2		BC_AAUHC_2		BC_BAUHG_1		BC_BAUHG_2		BC_CAUHG_1	
Code des paramètres		C	t	C	t	C	t	C	t	C	t
λ_y		.89	19.78 -2.39	.90	47.48 -5.10	.84	4.79 -.93	.87	19.33 -2.85	.78	5.32 -1.48
$\lambda_{x(1)}$		.89		.90		.64	1.23 -.69	.47	.86 -.97	-1.83	-2.25 -3.47
$\lambda_{x(2)}$		.89		.90		.64		.47		-1.83	
$\lambda_{x(3)}$		.89		.90		.64		.47		-1.83	
$\lambda_{x(4)}$		.89		.90		.64		.47		.79	1.15 -.31
$\lambda_{x(5)}$		.89		.90		.64		.47		-1.83	
ρ_1		.16	2.51	.18	2.77	.18	2.77	.17	2.59	.17	2.40
ρ_2		.21	2.92	.21	2.88	.20	2.88	.21	2.89	.19	2.71
ρ_{12}		.38	7.91	.37	6.77	.40	7.58	.40	7.69	.39	7.76
λ_{e1}		0.00		0.00		-.87	-1.75 -3.76	0.00		.03	.03 -1.16
λ_{e2}		13.02	1.32 1.22	0.00		0.00		12.18	1.29 1.18	0.00	
δ_1		0.00		0.00		2.77	.24	0.00		.52	.76
δ_2		.03	.26	2.00		0.00		.04	.28	0.00	
Code	Nom des variables	n	t	n	t	n	t	n	t	n	t
FADCB	Tarif réel du t.e.c.	.13	1.72	.12	1.92	.10	.69	.12	.78	.06	.49
PGRMD	Coût réel de l'essence	-.08	-.71	-.09	-.75	-.06	-.48	-.09	-.71	-.12	-.74
NCQR	Nombre d'automobiles	.55	2.64	.44	2.59	.55	2.29	.49	2.16	.11	.72
PCB	Ind. des prix à la cons.	-.14	-.65	-.02	-.10	-.10	-.48	.01	.04	.82	.49
RAWEQ	Salaire réel moyen	.81	2.73	.74	2.63	.81	.97	.81	.81	.69	.56
TW	Temps d'attente du t.e.c.	.55	3.14	.47	2.79	.50	1.42	.55	1.19	.53	.86
TT	Temps de trans. en t.e.c.	.04	.50	.04	.43	.01	.17	.05	.55	.05	.52
TAUTO	Temps de trans. en auto	-.84	-1.82	-.59	-1.61	-.73	-1.43	-.81	-1.72	-.54	-1.38
TLK	Inc. dû à la température	-.01	-.89	-.01	-.79	-.01	-.58	-.01	-.59	-.01	-.81
RFK	Inc. dû à la pluie	-.01	-1.53	-.01	-1.31	-.01	-.98	-.01	-1.45	-.00	-1.02
SFK	Inc. dû à la neige	.00	.56	.00	.63	.00	.69	.00	.62	.00	.68
CSFMR	Inc. dû à la neige acc.	-.02	-3.51	-.02	-3.38	-.02	-2.13	-.02	-3.48	-.01	-2.24
RCUTPN	Ind. pr. sc. (pr. & sec.)	-.02	-1.75	-.02	-1.69	-.02	-1.40	-.02	-1.77	-.02	-1.51
ISECON	Ind. pr. sc. (col. & un.)	-.02	-.95	-.01	-.59	-.01	-.61	-.02	-.84	-.01	-.62
RSIDCB	Inc. de dépla.-magasinage	.11	4.64	.10	4.23	.10	1.18	.12	1.02	.10	.80
EIQP	Niveau de l'emploi	.04	.46	.04	.41	.02	.16	.03	.27	.03	.26
EX7	Expo'67	.00	.24	.00	1.17	.00	.91	.00	1.19	.00	.97
STR	Grève à la C.T.C.U.M.	.00	1.62	.00	1.46	.00	1.43	.00	1.57	.00	1.24
VWM	Congé des travailleurs	.01	1.28	.01	1.97	.01	1.08	.01	1.06	.00	1.07
WD	Jours de travail/mois	.38	2.19	.38	2.23	.33	1.64	.38	1.71	.38	1.50
SAT	Samedis/mois	-.02	-.51	-.02	-.36	-.04	-.70	-.02	-.45	-.04	-.53
SHD	Dim. & jrs. de fêtes/mois	-.01	-.11	.00	.06	-.01	-.30	-.01	-.11	-.01	-.20
C	Constante	-.08	-.16	-.08	-.17	-.36	-.39	-.93	-.73	-50.89	-.68
$R^2(E(y))$		.980		.980		.980		.980		.980	
$R^2(E(y^{**}))$		.769		.925		.751		.753		.766	
$R^2(\hat{y}^{**})$		.769		.925		.751		.753		.766	
Variance résiduelle		.105		.089		.010		.096		.034	
$L(\omega^{**})$		178.13		171.53		179.07		178.51		180.12	

TABLEAU 14 (suite)

Valeurs des coefficients paramétriques, des élasticités, des tests statistiques et de $L(\omega^{**})$ pour divers modèles.

		$BC_D AUHG_1$		$BC_E AUHG_1$		$BC_{AA} AUHG_{1,2}$		$BC_{AA} AUHC_{1,2}$		$BC_{EE} AUHG_{1,2}$	
Code des paramètres		C	t	C	t	C	t	C	t	C	t
λ_y		.76	4.96 -1.53	.77	5.56 -1.69	.86	4.87 -1.81	1.50	82.97 27.74	.84	5.07 -.94
$\lambda_{x(1)}$		-1.99	-2.59 -3.89	.53	.45 -.39	.86		1.50		.50	.39 -.39
$\lambda_{x(2)}$		-1.99		1.55	.55 .19	.86		1.50		1.65	.57 .22
$\lambda_{x(3)}$		-1.99		-2.33	-1.92 -2.75	.86		1.50		-2.30	-1.74 -2.49
$\lambda_{x(4)}$		.72	1.08 -.43	.80	1.36 -.34	.86		1.50		.87	1.46 -.22
$\lambda_{x(5)}$		6.46	.76 .64	5.93	.88 .73	.86		1.50		5.46	.80 .65
ρ_1		.15	2.20	.18	2.62	.18	2.69	.22	3.15	.18	2.58
ρ_2		.18	2.80	.21	3.10	.20	2.91	.20	2.81	.21	2.97
ρ_{12}		.40	7.90	.33	6.40	.39	7.76	.34	5.78	.33	6.36
λ_{e1}		-.16	-.19 -1.36	-.26	-.27 -1.32	-.81	-1.27 -2.84	0.00		-.59	-.34 -.91
λ_{e2}		0.00		0.00		-1.69	-.05 -.09	0.00		15.97	.48 .45
θ_1		.65	.81	.64	.78	2.98	2.31	2.00		.85	.60
θ_2		0.00		0.00		-.26	-.37	2.00		.005	.07
Code	Nom des variables	n	t	n	t	n	t	n	t	n	t
FADCB	Tarif réel du t.e.c.	.06	.51	.11	.38	.11	1.16	.19	3.33	.12	.36
PGRMD	Coût réel de l'essence	-.13	-.79	-.01	-.13	-.05	-.44	.02	1.88	-.02	-.17
NCQR	Nombre d'automobiles	.10	.68	.70	1.85	.61	2.44	.50	3.19	.66	1.71
PCB	Ind. des prix à la cons.	.87	.53	-.28	-.39	-.19	-.90	.06	.34	-.23	-.36
RAWEQ	Salaire réel moyen	.66	.60	.67	.16	.80	2.36	.47	1.69	.67	.15
TW	Temps d'attente du t.e.c.	.51	.91	.72	.60	.50	2.91	.34	2.24	.69	.56
TT	Temps de trans. en t.e.c.	.05	.54	.00	.00	.01	.09	.06	.71	.03	.23
TAUTO	Temps de trans. en auto	-.51	-1.31	-.62	-.75	-.79	-1.63	-.35	-1.28	-.87	-1.51
TLK	Inc. dû à la température	-.01	-.72	-.01	-.77	-.01	-.73	-.06	-1.96	-.01	-.85
RFK	Inc. dû à la pluie	-.00	-.96	-.01	-1.25	-.01	-.98	-.03	-1.76	-.01	-1.27
SFK	Inc. dû à la neige	.00	.72	.00	.54	.00	.65	.01	.79	.00	.54
CSFMR	Inc. dû à la neige acc.	-.01	-2.09	-.01	-2.25	-.02	-2.12	-.04	-2.58	-.02	-2.02
RCUTPN	Ind. pr. sc. (pr. & sec.)	-.02	-1.42	-.01	-1.40	-.02	-1.34	-.13	-2.11	-.02	-1.31
ISECON	Ind. pr. sc. (col. & un.)	-.01	-.60	-.01	-.95	-.01	-.68	-.12	-1.48	-.02	-1.02
RSIDCB	Ind. de dépla.-magasinage	.11	.84	.11	.89	.10	3.73	.10	2.61	.12	.87
EIQP	Niveau de l'emploi	.02	.21	.04	.39	.03	.28	.13	1.49	.05	.50
EX7	Expo'67	.00	.94	.00	1.44	.00	.90	.00	.64	.00	1.37
STR	Grève à la C.T.C.U.M.	.00	1.27	.00	1.27	.00	1.48	.00	1.02	.00	1.17
VWM	Congé des travailleurs	.00	1.05	.00	1.07	.01	1.10	.04	2.99	.00	1.08
WD	Jours de travail/mois	.30	.16	.27	1.20	.32	1.78	.13	1.15	.26	.20
SAT	Samedis/mois	-.04	-.08	-.06	-.10	-.04	-.93	-.03	-1.02	-.05	-.10
SHD	Dim. & jrs. de fêtes/mois	-.02	-.07	-.04	-.08	-.02	-.40	-.03	-.79	-.04	-.08
C	Constante	-67.07	-.71	-7.78	-.72	.04	.08	1.59	2.69	-7.28	-.69
$R^2(E(y))$		.980		.981		.980		.978		.981	
$R^2(E(y^{**}))$		.770		.795		.765		.945		.770	
$R^2(\hat{y}^{**})$		.770		.795		.765		.945		.770	
Variance résiduelle		.030		.032		.009		.053		.039	
$L(\omega^{**})$		180.37		182.97		179.06		163.71		183.22	

La détermination de la forme fonctionnelle peut avoir des effets importants sur les estimations des paramètres des variables fixes et aléatoires des modèles. Une mauvaise spécification de la forme fonctionnelle et/ou du terme de l'erreur peuvent/peut amener l'introduction et/ou le rejet de variables explicatives et induire par conséquent de mauvais résultats[5].

A partir de nos exemples, nous allons voir si l'estimation conjointe de la forme fonctionnelle et des autres paramètres du modèle donne des résultats plus significatifs et plus raisonnables que ceux obtenus par des formes mathématiques choisies arbitrairement sur la partie fixe: ($\lambda_y = \lambda_{xk} = 1$, $\lambda_y = \lambda_{xk} = 0$), ou sur la partie aléatoire: ($X_m^{\delta_m=2}$), etc.

Premièrement, nous verrons dans quelle mesure une erreur de spécification sur la partie aléatoire du modèle associée à une forme mathématique contraignante pour une équation spécifique affecte le niveau de signification des variables. Deuxièmement, nous verrons dans quelle mesure une erreur de spécification du terme de l'erreur est critique pour le niveau de signification des variables si nous estimons les paramètres de la forme fonctionnelle. Et finalement, nous verrons si le relâchement graduel des contraintes sur la partie fixe a également un impact à ce niveau. Toutefois, avant de passer à une analyse de ces trois points, nous allons voir qu'un choix arbitraire de la forme fonctionnelle affecte le niveau de signification des variables.

Les statistiques-t de student qui vont servir d'éléments de comparaison entre les différents modèles choisis sont présentées dans le tableau 14.

5.- Une comparaison des modèles au niveau des tests statistiques ne peut cependant être très rigoureuse puisqu'une modification des hypothèses sur la partie aléatoire du modèle entraîne une variante dans la variance de l'erreur, donc de l'écart-type et par conséquent du student. Cependant, une comparaison entre les différentes hypothèses sur les parties fixe et stochastique du modèle s'impose puisqu'elle donnera une indication sur les difficultés qui peuvent survenir si la forme fonctionnelle ou/et si le terme de l'erreur est/sont mal spécifié(e), (s).

Une comparaison des modèles BC_*AU ($\lambda_y = \lambda_{xk} = 1$), BC_*Au ($\lambda_y = \lambda_{xk} = 0$) et BC_AAU ($\hat{\lambda}_y = \hat{\lambda}_{xk} = .62$) nous montre que les statistiques-t des différentes variables varient avec λ pour $\hat{\rho}_1 \neq \hat{\rho}_2 \neq \hat{\rho}_{12}$. Pour les premier et troisième modèles, les estimateurs associés à 10 variables sont significatifs à un niveau de confiance égal ou supérieur à 90%[6], contre seulement 5 pour le second. Ainsi, parmi les variables de prix, de revenu et de niveau de service, seule la variable du coût réel de l'essence PGRMD est significative dans le modèle logarithmique. Les variables associées au nombre d'automobiles NCQR, au revenu RAWEQ, au temps d'attente pour un mode concurrent TW et le temps de transport en automobile TAUTO sont, quant à elles, significatives dans les modèles linéaire et curvilinéaire. A ces variables, nous ajoutons également le prix du mode concurrent FADCB pour le modèle non-linéaire.

La considération des modèles BC_*, BC_*AU et BC_*AUHG_1, avec $\lambda_y = \lambda_{xk} = 1$, nous indique que les statistiques-t des variables varient avec ρ_ℓ, λ_{em} et δ_m. Le nombre de variables significatives à un niveau égal ou supérieur à .10 est respectivement égal à 14, 10 et 12. La valeur numérique de ces tests est, règle générale, plus élevée pour le modèle sans autocorrélation. (Dans le linéaire, les tests-t de student sont biaisés vers le haut si nous ne corrigeons pas pour l'autocorrélation des erreurs, ou si l'on veut, une correction de l'autocorrélation à partir d'un processus autorégressif d'un ordre multiple engendre une diminution dans la valeur du test-t. Nous avons donc une situation analogue à celle qui suggère que la correction de l'autocorrélation par un schéma autorégressif du premier ordre (si ρ_1 est positif) entraîne une diminu-

---

6.- Les points de pourcentage de la distribution t pour ∞ degrés de liberté (le nombre élevé de nos observations justifie ce degré de liberté) sont les suivants:

Niveau	.10	.05	.025	.01	.005
∞	1.282	1.645	1.960	2.326	2.576

tion dans la valeur numérique du test statistique. Une comparaison des modèles $BC_A AUHG_1$ et $BC_A HG_1$, $BC_A AU$ et BC_A (résultats non montrés pour BC_A) nous permet de conclure à une situation identique dans le non-linéaire). Si nous considérons à nouveau les huit premières variables de notre spécification, nous constatons que dans le premier modèle, seule la variable PGRMD n'est pas significative parmi les variables de prix et de revenu et la variable TT parmi les variables de service. Dans le modèle $BC_* AUHG_1$, la variable indice des prix à la consommation PCB est également exclue. Pour le modèle $BC_* AU$, au contraire et tel que déjà mentionné, seule la variable NCQR parmi les variables de prix est significative. La valeur statistique pour le test de student de chacune des huit premières variables est, pour le modèle $BC_* AUHG_1$, toujours inférieure à au moins une des deux autres.

La troisième approche nous suggère une comparaison des modèles $BC_A AU$ (10), $BC_A AUHG_1$ (9), $BC_A AUHC_1$ (14)[7], $BC_A AUHG_2$ (12), $BC_A AUHC_2$ (12), $BC_{AA} AUHG_{1,2}$ (9) et $BC_{AA} AUHC_{1,2}$ (15)[7]. (Les nombres entre parenthèses représentent le nombre de variables significatives au niveau .10.) Nous remarquons tout d'abord que les tests-t pour le modèle le plus général d'entre eux, à savoir $BC_{AA} AUHG_{1,2}$, demeurent identiques à ceux du modèle $BC_A AUHG_1$. La stabilité des coefficients estimés $\hat{\lambda}_y = \hat{\lambda}_{xk}$, $\hat{\rho}_1 \neq \hat{\rho}_2 \neq \hat{\rho}_{12}$, $\hat{\lambda}_{e1}$ et $\hat{\delta}_1$ entre les deux modèles explique cette réalité. Les paramètres $\hat{\lambda}_{e2}$ et $\hat{\delta}_2$ n'ont donc, dans le cas qui nous occupe, pratiquemement aucun impact sur les tests-t. Nous notons également que la variable PGRMD, laquelle n'est pas significative dans les modèles $BC_A AUHC_1$ et $BC_A AUHC_2$, devient significative à un niveau de 95% dans le modèle multiplicatif multivarié classique $BC_{AA} AUHC_{1,2}$. De plus, les valeurs des tests des variables de prix FADCB, PGRMD et NCQR sont plus élevées pour le modèle qui corrige avec la formulation classique de l'hétéroscédasticité. D'ailleurs,

---

7.- Comprend la constante.

ces deux premières variables de prix ne sont pas significatives dans le modèle[8] $BC_{AA}AUHG_{1,2}$. Le rejet des variables PCB et TT demeure, quant à lui, sans équivoque dans les sept modèles. Ces résultats nous montrent que la forme fonctionnelle de la partie fixe ainsi que les autres paramètres du modèle affectent le niveau de signification des variables; cependant, ils nous indiquent que les tests-t ne sont pas nécessairement les meilleurs là où, la valeur de la fonction de vraisemblance générée est la plus élevée.

L'approche suivante témoigne également de la dernière constatation. En effet, une comparaison entre les diverses applications du modèle BC_SAUHG_1 pour S = *, A, B, C, D, E, nous permet de constater que le nombre de variables indépendantes qui ont un niveau de signification supérieur à la valeur critique de 1.282 a une nette tendance à diminuer lorsque nous avons une plus grande famille paramétrique sur la partie fixe du modèle. Ce problème survient dans presque tous nos modèles quoique pour certains la diminution n'est pas assurée entre chacune des étapes. D'ailleurs pour le modèle précité, nous avons un hiatus à cette diminution causée par le modèle BC_EAUHG_1. Ainsi de 12 variables significatives à ce niveau dans la série *, nous passons à 9 dans la série A, à 7 dans B, à 4 dans C, à 3 dans D et à 5 dans E. De ces nombres, nous en comptons 5 parmi les variables de prix (PGRMD, NCQR), de revenu (RAWEQ) et de niveau de service (TW, TAUTO) pour BC_*AUHG_1 et 4 (NCQR, RAWEQ, TW, TAUTO) pour BC_AAUHG_1. Dans la série B, nous en comptons 3 (NCQR, TW, TAUTO), dans les séries C et D une chacune (TAUTO) et enfin 2 dans la série E (NCQR, TAUTO).

8.- En ce qui concerne la comparaison des valeurs statistiques-t de student entre les modèles qui incorporent une formulation classique ou une formulation générale de l'hétéroscédasticité, nous ne pouvons établir de consensus puisqu'elles sont parfois plus élevées dans les modèles où nous corrigeons avec la formulation classique [($BC_{AA}AUHC_{1,2}$: 14 tests-t sur 22 sont plus grands que dans $BC_{AA}AUHG_{1,2}$), (BC_AAUHC_1: 14 tests-t sont plus grands que dans BC_AAUHG_1)], et sont parfois plus élevées dans les modèles où nous corrigeons avec la formulation générale de l'hétéroscédasticité [(BC_AAUHG_2: 17 tests-t sont plus grands que dans BC_AAUHC_2)].

(Pour beaucoup de variables, comme par exemple RAWEQ [9] et TW, nous notons une baisse tangible de la valeur du test de student, si nous imposons une plus grande famille paramétrique.)

Une explication à la première constatation, laquelle nous a été proposée par M. Dagenais, est reliée aux lois conditionnelle et marginale de β.

Ainsi si nous avons par exemple $\hat{\beta}$, $\hat{\lambda}$ et si nous postulons les hypothèses suivantes: 1°) $\hat{\lambda}_{xk} = \hat{\lambda}_y$,

2°) $\hat{\lambda}_{xk} \neq \hat{\lambda}_y$,

nous avons alors:

1) la loi conditionnelle

$$\text{Var}\,(\hat{\beta} \mid \hat{\lambda}_{xk} = \hat{\lambda}_y) = E[\hat{\beta} - E(\hat{\beta} \mid \hat{\lambda}_{xk} = \hat{\lambda}_y)]^2, \text{ pour la 1° hypothèse,}$$

2) la loi marginale

$$\text{Var}\,(\hat{\beta}) = E[\hat{\beta} - E(\hat{\beta})]^2, \text{ pour la 2° hypothèse.}$$

Or, si la variance de $\hat{\beta}$ est plus grande que la variance de $\hat{\beta}$ étant donné $\hat{\lambda}_{xk} = \hat{\lambda}_y$, les écarts-types pour la seconde hypothèse émise seront biaisés vers le haut. Nos résultats tendent à le confirmer.

Les valeurs statistiques-t de student prennent des valeurs non raisonnables pour les paramètres λ_y dans un grand nombre de nos applications. L'explication sous-jacente à ce problème est attribuée au mauvais calcul de la matrice de covariance, tel que noté dans le chapitre II. De même, les valeurs calculées des tests-t pour les autres coefficients paramétriques et pour certaines variables explicatives sont surestimées.

9.- Les variables revenu réel RAWEQ et prix du mode concurrent FADCB, lesquelles sont significatives dans tous les modèles de la série A (sauf FADCB pour BC_AAUHG_1 et $BC_{AA}AUHG_{1\,2}$) deviennent non significatives dès que nous libérons la contrainte d'égalité entre λ_y et λ_{xk}. Le problème se pose également dans d'autres catégories de variables. La variable d'activité, ventes au détail RSIDCB en est un exemple.

* * *

La détermination de la forme fonctionnelle et la spécification de l'erreur peuvent également amener des signes erronés de paramètres importants.

Une analyse de chacune des variables nous montre que les signes affectant les paramètres estimés sont, dans notre cas, assez robustes entre chacun des modèles.

Dans la catégorie des variables de prix, seul le signe du coefficient du prix du transport en commun FADCB répond aux anticipations théoriques (signe positif). Le signe du coefficient du prix de l'essence PGRMD est positif pour les modèles sans autocorrélation BC_* ($\lambda_y = \lambda_{xk} = 1$), $BC_*(\lambda_y = \lambda_{xk} = 0)$, BC_AHG_1 et le modèle avec autocorrélation $BC_{AA}AUHG_{1,2}$ et va, par conséquent, dans la direction opposée à notre anticipation. En ce qui concerne la variable nombre d'automobiles NCQR[10], la seule modification survient si nous imposons une forme logarithmique et si nous corrigeons pour l'autocorrélation. Le signe négatif qui affecte alors l'estimateur de cette variable est non-attendu (la variable devient cependant non significative à un niveau .10). Outre le changement du signe de la variable NCQR, nous comptons sept autres modifications qui surviennent pour le modèle BC_*AU si nous préférons la forme logarithmique à la forme linéaire. Certains d'entre eux deviennent alors compatibles ou incompatibles avec la théorie économique. Le changement le plus intéressant est attribué à la variable revenu RAWEQ, laquelle implique que l'essence est un bien supérieur avec la forme linéaire et un bien inférieur avec la forme logarithmique. (Dans ce dernier cas, la variable n'est pas significative.) Alors si nous considérons que l'essence est un bien supérieur (le signe également positif dans chacun des autres modèles en atteste), le signe négatif qui

---

10.- Le nombre d'automobiles est considéré comme un prix dans le sens qu'un nombre accru de voitures en circulation réduit le coût marginal pour chaque usager de ce mode.

affecte le coefficient estimé de la variable PCB dans la plupart de nos modèles nous permet de conclure que l'essence est un bien complémentaire pour les biens autres que le transport. Dans d'autres cas ($BC_S AUHG_1$ pour S = C, D, $BC_B AUHG_2$, $BC_{AA} AUHG_{1,2}$), le signe positif nous permettrait de conclure à des biens substituts. (PCB n'est cependant pas significatif dans ces cas.) Nous pouvons également ajouter que le signe positif qui affecte FADCB, nous permet de conclure que la quantité d'essence vendue donc l'usage de l'automobile est un substitut au transport en commun.

Les variables de niveau de service à savoir le temps d'attente TW et le temps de transport TT (sauf dans $BC_* AU$ avec la forme linéaire) pour le mode de transport en commun présentent un signe (positif) qui va dans la direction anticipée. Quant à la variable temps de transport en automobile TAUTO, laquelle est considérée comme un effet de congestion, elle ne nous confère pas à priori l'attente d'un signe particulier. La densité de la circulation augmente la consommation d'essence par voyage et réduit simultanément la fréquence des voyages. Cependant, il est peut-être justifié d'anticiper un plus grand effet sur la consommation accrue d'essence. Nos résultats ne justifient pas nos anticipations puisque nous obtenons, sauf pour le modèle $BC_* AU$ avec la forme logarithmique, un signe négatif. Il est notable que les variables TT et TAUTO pour le modèle $BC_* AU$ ($\lambda_y = \lambda_{xk} = 0$) ont le signe attendu contrairement au modèle $BC_* AU$ ($\lambda_y = \lambda_{xk} = 1$).

Si nous considérons la totalité de nos modèles, nous notons à postiori que les variables de confort, à savoir les températures chaudes et froides TLK, la pluie RFK ainsi que la neige accumulée au sol CSFMR, diminuent la part modale de l'automobile et ont donc un impact négatif sur la quantité d'essence vendue. Au contraire la neige SFK attire les clients potentiels de ce mode.

Pour les variables d'activité et d'agrégation, les variantes de la

forme fonctionnelle et de la spécification de l'erreur n'entraînent pas de modification substantielle puisque le signe affectant les paramètres estimés est maintenu entre chacun des modèles pour une même variable à quelques exceptions près. Les variables inscriptions aux collèges et universités RCUTPN, aux écoles primaires et secondaires ISECON,[11] ainsi que les variables d'agrégat, nombre de jours mensuels de samedis SAT, de dimanches et jours fériés SHD présentent un signe négatif, lequel est contraire à nos anticipations, sauf pour les modèles BC_*AU ($\lambda_y = \lambda_{xk} = 0$) et BC_* ($\lambda_y = \lambda_{xk} = 0$). Un signe positif affecte les variables ISECON et SHD dans le premier modèle et les variables RCUTPN, WD et SAT dans le second. Quant aux autres caractéristiques présentent dans ces deux catégories de variables, elles exhibent toujours le signe positif attendu, sauf EIQP dans le modèle BC_* ($\lambda_y = \lambda_{xk} = 0$).

La forme fonctionnelle n'affecte pas le signe des coefficients autorégressifs, lequel est positif et par conséquent conforme à nos attentes.

En résumé, nous pouvons dire que la forme fonctionnelle et la spécification de l'erreur affectent les signes des coefficients des régresseurs, mais dans notre cas, aucun des modèles ne se distingue vraiment des autres en présentant les meilleurs signes, si ce n'est le modèle BC_*AU ($\lambda_y = \lambda_{xk} = 0$).

Les mauvais signes associés à certaines variables peuvent être expliqués sans doute par la multicollinéarité entre celles-ci. Ainsi, dans le modèle $BC_{AA}AUHG_{1,2}$, la corrélation partielle entre TAUTO et NCQR s'élève à .89, entre RCUTPN et ISECON à .65 et entre PCB [FADCB (-.71), PGRMD (-.68), NCQR (.53), RAWEQ (.93), TW (.81), RSIDCB (-.57), WD (.50)], etc.. Dans le modèle

---

11.- Le signe négatif affectant les variables d'activités RCUTPN et ISECON peut être expliqué par l'absence d'une variable d'activité importante dans la liste des régresseurs, à savoir une variable d'activité sociale (déplacements pour motifs divers). Nous anticipons que cette dernière variable serait reliée négativement avec les variables RCUTPN et ISECON.

$BC_{AA}AUHC_{1,2}$, le mauvais signe de PGRMD est peut-être également expliqué par une corrélation avec FADCB (.99), PCB (-.94), RAWEQ (-.95), TW (-.95), TT (-.94) WD (-.92), SAT (-.80) et autres.

Le tableau 15 présente le nombre de paires distinctes de corrélation partielle supérieure ou égale à .50 pour différents niveaux α.

TABLEAU 15

Nombre de paires de coefficients de corrélation partielle pour différents niveaux α.

Modèles	Niveaux					
	.90	.80	.70	.60	.50	Total
BC_{*} (λ=1)	6	5	11	16	12	50
$BC_{*}AU$ (λ=1)	3	3	1	7	8	22
$BC_{*}AU$ (λ=0)	0	0	1	0	4	5
$BC_{*}AUHG_{1}$ (λ=1)	2	9	10	7	15	43
BC_{A}	6	5	10	17	12	50
$BC_{A}AU$	3	2	2	4	8	19
$BC_{A}HG_{1}$	10	17	18	21	13	79
$BC_{A}HC_{1}$	10	10	11	16	21	68
$BC_{A}AUHG_{1}$	2	5	9	8	10	34
$BC_{A}AUHC_{1}$	16	7	15	11	11	60
$BC_{A}AUHG_{2}$	0	4	1	5	12	22
$BC_{A}AUHC_{2}$	12	14	11	13	12	62
$BC_{B}AUHG_{1}$	2	6	9	8	7	32
$BC_{B}AUHG_{2}$	1	4	4	4	11	24
$BC_{C}AUHG_{1}$	15	11	5	7	7	45
$BC_{D}AUHG_{1}$	6	4	3	3	11	27
$BC_{E}AUHG_{1}$	1	10	5	4	5	25
$BC_{AA}AUHG_{1,2}$	2	6	8	8	13	37
$BC_{AA}AUHC_{1,2}$	14	12	10	20	10	66
$BC_{EE}AUHG_{1,2}$	3	4	6	6	8	27

Pour le modèle BC_* par exemple, la matrice de corrélation entre les variables indépendantes montre que sur 231 paires de corrélation partielle r_{ij} où

$$r_{ij} = \frac{\Sigma (X_i^{**} - \overline{X}_i)(X_j^{**} - \overline{X}_j)}{\sqrt{\Sigma (X_i^{**} - \overline{X}_i)^2 \Sigma (X_j^{**} - \overline{X}_j)^2}} ,$$

50 paires de régresseurs ont des coefficients de corrélation supérieurs à .50 dont 6 ont des coefficients plus grands que .90 et 5 plus grands que .80.

Il est intéressant de noter que:

1°) la correction pour l'autocorrélation diminue la colinéarité dans le linéaire et dans le non-linéaire,

2°) la correction pour l'hétéroscédasticité avec la formulation générale en introduit,

3°) la correction pour l'hétéroscédasticité avec la formulation classique en introduit beaucoup plus. (Ceci est vrai pour les trois applications présentées si nous corrigeons de plus pour l'autocorrélation. Dans le cas où nous n'introduisons pas de processus autorégressif, la situation inverse est constatée.)

Nos constatations sont basées sur les seuls résultats présentés dans ce rapport. Aussi, certains autres auraient peut-être suggéré des affirmations opposées.

4.3.3. Effets de la forme fonctionnelle et de la spécification de l'erreur sur les élasticités.

Dans les paragraphes suivants, nous allons voir sur la base du concept d'élasticité, que la forme fonctionnelle et que la spécification de l'erreur modifient l'influence apparente des diverses caractéristiques économiques et autres.

La variation dans les élasticités calculées pour les différents modèles est "d'un point de vue appliqué, probablement le critère le plus crucial pour distinguer les propriétés d'un modèle à celles d'un autre", (Gaudry et Wills, 1978, p.8).

Une modification des élasticités dans les variables de prix peut être très importante par exemple, pour les preneurs de décision au niveau des politiques gouvernementales, etc..

Dans le contexte d'un modèle de demande, l'élasticité de y par rapport à toutes les variables X_k est définie comme étant:

$$\eta_k = \frac{\partial y}{\partial X_k} \cdot \frac{X_k}{y} \quad . \tag{54}$$

Si $\lambda_y \neq 0$ et si $\lambda_{xk} \neq 0$, la dérivée analytique est:

$$\eta_k = \frac{\partial y}{\partial y^{(\lambda_y)}} \cdot \frac{\partial y^{(\lambda_y)}}{\partial X_k^{(\lambda_{xk})}} \frac{\partial X_k^{(\lambda_{xk})}}{\partial X_k} \cdot \frac{X_k}{y}$$

$$\eta_k = \beta_k (X_k)^{\lambda_{xk}-1} X_k [(y)^{\lambda_y - 1} y]^{-1} = \beta_k X_k^{\lambda_{xk}} y^{-\lambda_y} \quad . \tag{55}$$

Ce concept d'élasticité est un cas spécial des élasticités dérivées dans le texte de Gaudry et Wills, (1978, pp.43-44).

L'élasticité de la demande est calculée aux valeurs moyennes de y et de X_k ($\bar{y}$ et $\bar{X}_k$). Par conséquent,

$$\bar{\eta}_k = \beta_k \bar{X}_k^{\lambda_{xk}} \bar{y}^{-\lambda_y} \quad . \tag{56}$$

$$\text{Si } \lambda_y = \lambda_{xk} = 1, \quad \bar{\eta}_k = \beta_k \bar{X}_k \bar{y}^{-1} \tag{57}$$

et

$$\text{si } \lambda_y = \lambda_{xk} = 0, \quad \bar{\eta}_k = \beta_k \quad . \tag{58}$$

Dans ce dernier cas, nous avons le concept "d'élasticité constante"; (les élasticités ne sont pas dépendantes du niveau des variables). Dans les autres cas, la mesure d'élasticité varie de mois en mois.

En première analyse, nous remarquons les élasticités très élevées de la constante dans les modèles BC_CAUHG_1 (-50.89) et BC_DAUHG_1 (-67.07). Nous expliquons ces élasticités exagérées par les valeurs négatives de certains paramètres affectant les variables transformables X_k. Dans le second modèle mentionné ci-dessus, un coefficient de -1.99 a, au contraire d'un effet de puissance, un effet qui tend à réduire les variables de prix, de revenu et de niveau de service sous la forme d'une constante. La transformation Box-Cox élimine ainsi des variables importantes du modèle et la constante va capter toute l'influence.

En ce qui concerne les différentes variables, nous limiterons notre analyse principalement aux huit premières. Outre une modification légère de l'élasticité de la demande par rapport à la variable EIQP (.23) et à la variable ISECON (-.19) dans le modèle BC_AHG_1 (l'élasticité de la demande par rapport à ces deux variables est dans chacun des autres cas, inférieure à .10), ainsi que par rapport à la variable RSIDCB ($\eta_k \simeq .10$) et à la variable WD (η_k varie entre .13 et .56), il n'y a pas de remarques intéressantes à formuler sur ces variables, puisqu'elles ne concourrent que très peu à l'explication de la demande et ce, dans chacun des modèles.

Si nous faisons des comparaisons sur la base des différentes applications, nous notons que les ordres de grandeur de certaines élasticités sont modifiés et que l'importance stratégique des différents facteurs diffère quelque peu si la forme fonctionnelle et la spécification de la partie stochastique

sont modifiées.

Si nous considérons l'importance stratégique des variables qui ont un plus fort pouvoir explicatif de la consommation d'essence, nous notons sur la base des coefficients statistiquement significatifs que la variable revenu RAWEQ est celle qui, règle générale, influence le plus la consommation de ce bien. Ainsi, dans le modèle BC_AAUHG_1 ($\hat{\lambda}_y = \hat{\lambda}_{xk} = .87$; $\hat{\rho}_\ell$; $\hat{\lambda}_{el} = -.83$; $\hat{\delta}_1 = 2.88$), la variable revenu est celle qui contribue le plus à l'explication de la demande, suivie de près par le temps de transport TAUTO et puis par le nombre d'automobi NCQR. Dans le modèle BC_AAUHC_1 ($\hat{\lambda}_y = \hat{\lambda}_{xk} = 1.27$; $\hat{\rho}_\ell$; $\lambda_{el} = 0$; $\delta_1 = 2$), le revenu et le nombre d'automobiles expliquent similairement le choix du mode de transport en automobile, suivis par la variable temps de transport en automobile. Il est intéressant de noter que si nous libérons la contrainte d'égalité entre les paramètres λ_{xk} (modèle BC_EAUHG_1), la variable temps de transport en automobile contribue le plus à expliquer la demande d'essence. Si nous imposons la forme logarithmique (avec correction pour l'autocorrélation des résidus), la variable coût réel de l'essence PGRMD devient (toujours sur la base des variables explicatives) la plus explicative.

Nous allons maintenant examiner les ordres de grandeur des élasticités

L'élasticité-prix croisée de la demande par rapport au tarif du transport en commun FADCB semble assez raisonnable. La modification la plus sensible survient si nous imposons la forme logarithmique et si nous corrigeons pour l'autocorrélation. Cette élasticité passe alors de .18 ($BC_\star$) à .01 ($BC_\star AU$). C'est dans ce dernier modèle que les élasticités apparaissent être les plus basses. Dans chacune des applications, l'essence demeure un bien très inélastique par rapport au prix du mode concurrent, l'élasticité variant entre .06 et .19. Par conséquent, une hausse des tarifs du transport en commun ne provoque pas une baisse très marquée des usagers en faveur d'un mode concurrent,

l'automobile. Nous avons mentionné antérieurement que l'automobile est un substitut pour le transport en commun. Nous pouvons ajouter ici que c'est un substitut assez faible. Dans le cas où l'élasticité-tarif est égale à .01, le coefficient du revenu est négatif et nous ne pouvons alors conclure, à partir de ce modèle, à la substitution entre les deux modes.

L'élasticité-prix propre de la demande d'essence PGRMD varie entre -.01 et -.16 et est donc très basse. Cette forte inélasticité de la demande par rapport au coût de l'essence suggère donc que les producteurs de pétrole pourraient augmenter impunément le prix de ce produit sans subir une baisse tangible dans les quantités vendues, si aucun contrôle n'est exercé ou s'il y a constitution de monopoles ou de cartels. Ceci n'est pas tellement rassurant puisqu'une hausse des prix n'entraînerait pas la résorbation de la crise de l'énergie.

L'élasticité de la demande par rapport au nombre d'automobiles n'est pas très inélastique (sauf pour les modèles BC_CAUHG_1 ($\eta_{nbr.d'autos}$=.11)[12] et BC_DAUHG_1 ($\eta_{nbr.d'autos}$=.10)), ce qui est très raisonnable. Elle varie entre .44 dans le modèle BC_AAUHG_2 et .70 dans le modèle BC_EAUHG_1. Ainsi, une hausse de 10% dans le nombre d'automobiles impliquerait une augmentation de 4.4% dans la vente d'essence pour le premier modèle et un accroissement de 7% pour le second.

En ce qui a trait à l'élasticité de la demande de l'essence par rapport aux biens autres que le transport (élasticité croisée), elle est également très inélastique sauf dans les deux modèles où l'élasticité de la demande par

---

12.- Cette variable qui n'est pas significative dans les modèles BC_CAUHG_1 et BC_DAUHG_1, l'est fortement dans les modèles BC_CAUHC_1 et BC_DAUHC_1. Dans ces deux derniers cas, $\eta_{nbr.d'autos}$ est respectivement égale à .61 et .62 (résultats non montrés).

rapport au nombre d'automobiles est justement très inélastique. Dans ces deux modèles (BC_CAUHG_1 et BC_DAUHG_1), l'élasticité-prix est respectivement égale à .82 et .87. (Associés au signe positif de la variable revenu, ces deux derniers résultats suggéreraient, si PCB était significatif, que l'essence est un fort substitut pour les biens autres que le transport[13].)

En ce qui concerne l'élasticité de la variable revenu, il y a très peu à dire si ce n'est qu'elle nous permet d'affirmer que l'essence est un bien supérieur nécessaire ($\eta_{revenu}<1$). Certains modèles nous permettent même de conclure à un bien très nécessaire; dans le modèle BC_AAU, $\eta_{revenu}=.82$.

Pour les variables de niveau de service, il est intéressant de noter qu'il existe une assez grande stabilité de la demande d'essence par rapport au temps de transport en autobus (l'élasticité varie entre .01 et .08, ce qui est relativement bas), contrairement aux deux autres variables impliquées dans cette catégorie. Ainsi, le passage du modèle BC_* ($\lambda_y = \lambda_{xk} = 0$) au modèle BC_*AU ($\lambda_y = \lambda_{xk} = 0$) nous permet de noter une diminution de l'élasticité dérivée par rapport au temps d'attente du transport en commun de .30 à .04. Par ailleurs, la libération de la contrainte de fixité des λ_{xk}, ainsi que la correction pour l'autocorrélation et la correction pour l'hétéroscédasticité avec la formulation générale si *RSIDCB* est choisi pour X_m dans $f(X_t)$, entraîne un accroissement jusqu'à un ordre de grandeur de .70. L'élasticité du temps d'attente TW est plus élevée que celle du temps de transport TT sauf dans le modèle BC_*AU (forme logarithmique). Les gens sont donc plus sensibles au temps d'at-

---

13.- Avec la formulation classique de l'hétéroscédasticité, ces élasticités sont égales (pour les modèles BC_SAUHC_1, pour S = C, D) à -.15 et à -.16 respectivement (résultats non montrés). Nous aurions alors des biens complémentaires. Nous pouvons également ajouter que pour les séries C et D, cette variable a toujours un signe négatif si nous ne corrigeons pas pour l'autocorrélation (BC_C: $\eta_{prix}=-.25$) et positif si nous corrigeons (BC_CAU: $\eta_{prix}=.74$) excepté si nous corrigeons également pour l'hétéroscédasticité avec la formulation classique, tel que noté ci-dessus.

tente qu'au temps de transport, ce à quoi on s'attendait à priori. Cependant, le rapport TW/TT donne dans la plupart de nos modèles des résultats aberrants. Ce rapport pour quelques-uns de nos modèles est présenté dans le tableau 16.

TABLEAU 16

Rapport des élasticités du temps d'attente sur le temps de transport pour quelques modèles, (en valeur absolue).

BC_* (λ=1)	BC_AAU	BC_AHG_1	BC_AAUHG_1	$BC_{EE}AUHG_{1,2}$
6	53	5.5	50	23

Ainsi, une variation de une minute dans le temps d'attente a un impact de 5.5 fois plus élevé que celui du temps de transport sur la demande d'essence dans le modèle BC_AHG_1 et de 50 fois dans le modèle où nous corrigeons pour l'autocorrélation.

La variable temps de transport en automobile TAUTO est très sensible à la forme fonctionnelle et à la spécification du terme de l'erreur. Dans le modèle BC_AAU, l'élasticité propre de la demande par rapport à cette variable s'élève à -.83. Dans les modèles BC_AHG_1 et BC_AAUHG_1, elle s'établit respectivement à -.06 et à -.78. Dans chacun des autres modèles, elle n'est pas très inélastique, sauf à quelques occasions, ce qui n'est pas raisonnable, d'autant plus que le signe qui l'affecte est négatif. Ceci suggère donc qu'une augmentation de 100% du temps de transport en automobile, réduit énormément la fréquence des voyages, et par conséquent la quantité vendue de l'essence. Pour les automobilistes, nous constatons que dans 17 cas sur 19, ils sont plus sensibles au temps de trajet qu'au prix. Ces résultats confirment nos attentes, puisque les automobilistes choisissent habituellement ce mode de transport pour diminuer leur temps de déplacement, parce que les distances à franchir sont grandes, parce qu'il n'y a pas, pour toutes les villes de la province de

Québec, un réseau de transport bien structuré et par conséquent un service bien établi, etc.. Encore ici cependant, le rapport est dans beaucoup de cas exagéré, tel que montré dans le tableau 17.

TABLEAU 17

Rapport des élasticités du temps de trajet en automobile sur le prix de l'essence pour quelques modèles (en valeur absolue).

BC_* (λ=1)	$BC_A AU$	$BC_A HG_1$	$BC_A AUHG_1$	$BC_D AUHG_1$	$BC_E AUHG_1$	$BC_{EE} AUHG_{1,2}$
2.40	13.83	.43	15.60	3.92	92.00	43.50

De plus, l'élasticité du temps de transport en automobile est, règle générale, supérieure à celle du temps d'attente quoique dans certains modèles l'inverse se produit (v.g. tableau 18). Nous croyons que les résultats sont assez raisonnables, si nous considérons que le temps de transport en automobile [une fois que les gens sont captifs de ce mode] influence davantage la consommation d'essence.

TABLEAU 18

Rapport des élasticités du temps d'attente sur le temps de trajet en automobile pour quelques modèles (en valeur absolue).

BC_* (λ=1)	$BC_A AU$	$BC_A HG_1$	$BC_A AUHG_1$	$BC_D AUHG_1$	$BC_{EE} AUHG_{1,2}$
2.50	.63	8.67	.64	1.00	.79

Les tableaux 19 et 20 présentent les prix implicites du temps[14]. Ainsi par exemple, en consultant les valeurs inscrites dans le tableau 19, nous

14.- Les prix du temps sont obtenus par le rapport suivant:

$$\frac{\eta_t}{\eta_p} \cdot \frac{\bar{p}}{\bar{t}} \cdot 60$$

pouvons obtenir quelle variation dans le coût réel de l'essence impliquerait la même perte du nombre d'automobilistes qu'une variation d'une heure dans chacune des variables de temps.

TABLEAU 19

Valeur implicite des prix du temps pour quelques modèles ($/heure).
- $\bar{p}$ ≡ PGRMD -

Modèles	Prix du temps d'attente	Prix du temps de trajet en autobus	Prix du temps de trajet en auto
$BC_{\star}$ (λ=1)	31.71	1.06	5.13
$BC_A AU$	46.70	.18	29.55
$BC_A HG_1$	8.31	.30	.92
$BC_A AUHG_1$	52.86	.21	33.33
$BC_A AUHG_2$	36.34	.53	22.44
$BC_A AUHC_2$	27.60	.47	14.00
$BC_C AUHG_1$	23.35	.44	9.62
$BC_D AUHG_1$	20.74	.41	8.38
$BC_{EE} AUHG_{1,2}$	182.36	1.59	92.95

Dans le modèle $BC_{\star}$, ceci signifie que les gens seraient prêts à payer $ 31.71 de plus l'essence pour attendre une heure de moins, à payer $ 5.13 de plus pour diminuer le temps de leur déplacement en automobile et $ 1.06 celui en autobus. Dans chacun de ces modèles, le prix du temps de transport en automobile est supérieur à celui en autobus ce qui est fort surprenant. Par ail-

---

où

η_t: représente le taux d'élasticité propre et croisée de la demande par rapport au temps de trajet et l'élasticité croisée de la demande par rapport au temps d'attente,

η_p: représente le taux d'élasticité propre (croisée) par rapport au prix-coût de l'essence (tarif),

$\bar{p}$, $\bar{t}$: sont respectivement les valeurs moyennes du prix et des temps de trajet et d'attente (TAUTO, TT, TW).

leurs dans chacun d'eux, le prix du temps d'attente est celui qui est le plus élevé, ce qui est raisonnable, quoique du point de vue de leur valeur numérique, les résultats sont beaucoup trop élevés. Parmi toutes ces variantes, le prix du temps d'attente qui semble le plus raisonnable est celui obtenu dans le modèle BC_AHG_1. Le prix du temps de trajet en auto apparaît être le plus raisonnable dans le modèle BC_*.

Dans le tableau 20, nous avons calculé quelle variation dans le prix du mode concurrent entraînerait la même perte du nombre d'automobilistes qu'une modification d'une heure dans chacune des variables de temps.

TABLEAU 20

Valeur implicite des prix du temps pour quelques modèles ($/heure).
- p ≡ FADCB -

Modèles	Prix du temps d'attente	Prix du temps de trajet en autobus	Prix du temps de trajet en auto
BC_* (λ=1)	6.62	.22	1.07
BC_AAU	12.87	.05	8.15
BC_AHG_1	4.11	.15	.45
BC_AAUHG_1	11.04	.04	6.96
BC_AAUHG_2	10.27	.15	6.34
BC_AAUHC_2	9.51	.16	4.83
BC_CAUHG_1	21.45	.41	8.84
BC_DAUHG_1	20.64	.41	8.34
$BC_{EE}AUHG_{1,2}$	13.96	.12	7.12

Une comparaison des tableaux 19 et 20 semble suggérer que les individus sont moins sensibles à une variation dans le coût réel de l'essence qu'à une augmentation dans le tarif - sauf dans le modèle BC_DAUHG_1 -. Par conséquent, les gens seraient prêts à subir une plus grande augmentation dans le coût réel de l'essence que dans celle du tarif du transport en commun ($ 8.31

contre $ 4.11 dans le modèle BC_AHG_1) pour attendre par exemple une heure de moins.

Il n'y a rien de fondamental à formuler sur les valeurs numériques des paramètres autorégressifs, si ce n'est que la valeur de ρ_{12} est beaucoup plus élevée lorsque λ_y est égal à zéro. La valeur inférieure de ρ_1 dans chacun des modèles implique que nous sommes loin d'un modèle en différences premières et qu'il aurait été inapproprié de formuler un tel modèle.

L'impact de la forme fonctionnelle et des paramètres sur la valeur des élasticités est explicitement exhibé par nos résultats. Outre ce qui a déjà été mentionné, nous pouvons ajouter que nous obtenons dans notre cas, des élasticités inférieures si nous préférons la formulation classique de l'hétéroscédasticité à la formulation générale et ce, quelle que soit la forme fonctionnelle de la partie fixe.

Les élasticités font ressortir des particularités différentes et amènent à des conclusions différentes. Cependant, même si nous pouvons considérer que certains modèles semblent présenter des taux d'élasticités plus compatibles à nos anticipations ou du moins plus raisonnables, nous ne pouvons dire qu'un modèle présente les "vraies" valeurs.

CONCLUSION

La complexité de la recherche dans le domaine de la non-linéarité nous a amené à formuler de nombreuses questions auxquelles certaines explications ont parfois été avancées, bien que celles-ci ne soient pas formelles et qu'elles demeurent par conséquent toujours sujettes à discussion. Néanmoins, les résultats que nous avons obtenus justifient l'ophélimité (ou satisfaction) que nous retirons d'un travail aussi complexe et ils peuvent être utilisés comme apologétique de l'utilité de la détermination conjointe de la forme fonctionnelle de la partie fixe et de la partie stochastique du modèle.

Nos résultats montrent que la transformation Box-Cox sur la partie fixe peut changer le caractère hétéroscédastique des résidus (ou compenser sans corriger complètement, ou en introduire) et que par conséquent l'hétéroscédasticité, tout comme la non-linéarité, est une force qui agit sur la forme fonctionnelle. L'introduction de la forme générale de l'hétéroscédasticité permet d'établir le discernement entre ces deux forces qui agissent sur la forme fonctionnelle si nous ne corrigeons pas avec la vraie forme de l'hétéroscédasticité.

Ils montrent également que l'estimation de la forme fonctionnelle optimale, laquelle est estimée conjointement avec les autres paramètres du modèle, affecte les élasticités dérivées de la demande, les tests statistiques, etc. Ils suggèrent donc que les élasticités obtenues d'une forme fonctionnelle choisie arbitrairement et/ou d'une matrice variance-covariance dont les éléments sont mal spécifiés (absence d'un processus autorégressif, formulation arbitraire de l'hétéroscédasticité) peuvent être considérées comme non orthodoxes et que l'utilisation d'une forme fonctionnelle inexacte et/ou d'une mauvaise spécification du terme de l'erreur peut engendrer des erreurs de spécification

de nature indéterminée.

Sur la base du seul concept de la vraisemblance de l'échantillon, il nous a été permis de rejeter d'une manière significative de nombreuses applications, comme par exemple les modèles dans lesquels nous utilisons une formulation classique de l'hétéroscédasticité par opposition à la formulation générale. Sur la base du concept d'élasticité, certains modèles se sont dissociés des autres en présentant des résultats plus raisonnables, mais ceux-ci ne sont pas associés au modèle qui optimise la fonction de vraisemblance.

APPENDICE 1

LOGARITHME DE LA FONCTION DE VRAISEMBLANCE CONCENTREE.

Si nous substituons l'expression (37) dans (36), nous pouvons réécrire la fonction de vraisemblance comme étant égale à:

$$\mathcal{L}(y;\ \beta_k,\ \sigma^{**2},\ \rho_\ell,\ \lambda_y,\ \lambda_{xk},\ \lambda_{em},\ \delta_m) = \prod_{t=1+r}^{n} \frac{y_t^{\lambda_y - 1}}{\sqrt{2\pi}\,\sigma^{**}\sqrt{\exp\left(\sum_m \delta_m X_{m_t}^{(\lambda_{em})}\right)}} \cdot$$

$$\exp\left\{ -\frac{1}{2\sigma^{**2}} \left[\left(\frac{y_t^{(\lambda_y)}}{\sqrt{\exp\left(\sum_m \delta_m X_{m_t}^{(\lambda_{em})}\right)}} - \sum_{\ell=1}^{r} \rho_\ell \frac{y_{t-\ell}^{(\lambda_y)}}{\sqrt{\exp\left(\sum_m \delta_m X_{m_{t-\ell}}^{(\lambda_{em})}\right)}} \right) - \sum_k \beta_k \right.\right.$$

$$\left.\left. \left(\frac{X_{k_t}^{(\lambda_{xk})}}{\sqrt{\exp\left(\sum_m \delta_m X_{m_t}^{(\lambda_{em})}\right)}} - \sum_{\ell=1}^{r} \rho_\ell \frac{X_{k_{t-\ell}}^{(\lambda_{xk})}}{\sqrt{\exp\left(\sum_m \delta_m X_{m_{t-\ell}}^{(\lambda_{em})}\right)}} \right) \right]^2 \right\} =$$

$$\prod_{t=1+r}^{n} \frac{y_t^{\lambda_y - 1}}{\sqrt{2\pi}\,\sigma^{**}\sqrt{\exp\left(\sum_m \delta_m X_{m_t}^{(\lambda_{em})}\right)}} \cdot \exp\left\{ -\frac{1}{2\sigma^{**2}} \left[y_t^{**} - \sum_k \beta_k X_{k_t}^{**} \right]^2 \right\} \quad .(1.1)$$

a) Le logarithme de la fonction de vraisemblance non concentrée.

Le logarithme de la fonction de vraisemblance non concentrée correspondant à l'expression (1.1) est:

$$L(y;\ \beta_k,\ \sigma^{**2},\ \rho_\ell,\ \lambda_y,\ \lambda_{xk},\ \lambda_{em},\ \delta_m) =$$

$$L(\omega) = \sum_{t=1+r}^{n} \left\{ (\lambda_y - 1)\ \text{ln}\ y_t - \frac{1}{2}\ \text{ln}\ \sigma^{**2} - \frac{1}{2}\ \text{ln}\ f(X_t) \right.$$

$$- \frac{1}{2\sigma^{**2}} [y_t^{**} - \Sigma\beta_k X_{k_t}^{**}]^2 \Bigg\} \qquad (1.2)$$

$$= (\lambda_y - 1) \sum_{t=1+r}^{n} \ell n\, y_t - \left(\frac{n-r}{2}\right) \ell n\, \sigma^{**2} - \frac{1}{2} \sum_{t=1+r}^{n} \ell n\, f(X_t)$$

$$- \frac{1}{2\sigma^{**2}} \sum_{t=1+r}^{n} [y_t^{**} - \Sigma\beta_k X_{k_t}^{**}]^2 \qquad (1.3)$$

où $\left(\frac{n-r}{2}\right)$ provient de ce que $\frac{1}{2} \ell n\, \sigma^{**2}$ est défini comme une constante pour une somme de $(1 + r)$ jusqu'à n. Le terme constant (π) n'apparaît pas dans la fonction puisqu'il n'intervient pas dans le processus de maximisation (Tran Cong Liem, 1979, p.2).

b) Le logarithme de la fonction de vraisemblance concentrée sur σ^{**2} et sur les β_k.

Si nous maximisons (1.3) par rapport à σ^{**2} nous obtenons,

$$\frac{\partial L(\omega)}{\partial \sigma^{**2}} = - \left(\frac{n-r}{2}\right) \frac{1}{\sigma^{**2}} + \frac{1}{2(\sigma^{**2})^2} \sum_{t=1+r}^{n} [y_t^{**} - \Sigma\beta_k X_{k_t}^{**}]^2 \quad . \qquad (1.4)$$

En posant la dérivée égale à zéro, nous obtenons:

$$(\hat{\sigma}^{**2}) = \sum_{t=1+r}^{n} \frac{[y_t^{**} - \Sigma\beta_k X_{k_t}^{**}]^2}{(n-r)} = \frac{1}{(n-r)} e'_t e_t \qquad (1.5)$$

et en remplaçant σ^{**2} dans (1.3) par $\hat{\sigma}^{**2}$, nous avons le logarithme de la fonction de vraisemblance concentrée sur σ^{**2}. Ainsi,

$$L(y;\, \beta_k,\, \rho_\ell,\, \lambda_y,\, \lambda_{xk},\, \lambda_{em},\, \delta_m) =$$

$$L(\omega^*) = (\lambda_y - 1) \sum_{t=1+r}^{n} \ell n\, y_t - \left(\frac{n-r}{2}\right) \ell n\, \hat{\sigma}^{**2} - \frac{1}{2} \sum_{t=1+r}^{n} \ell n\, f(X_t) \quad . \qquad (1.6)$$

Suite à une substitution de (1.5) dans (1.3), le dernier terme de cette dernière expression devient équivalent à une constante, à savoir - (n-r/2), et peut par conséquent être éliminé.

Pour concentrer la fonction (1.6) sur les β_k, nous devons présenter l'expression (1.5) sous une forme matricielle. Ainsi, la dérivée première de (1.6) par rapport au vecteur β est égale à:

$$\frac{\partial L(\omega^*)}{\partial \beta} = - \left(\frac{n-r}{2}\right) \frac{1}{\hat{\sigma}^{**2}} \cdot \frac{\partial \hat{\sigma}^{**2}}{\partial \beta} \qquad (1.7)$$

avec

$$\hat{\sigma}^{**2} = \frac{1}{(n-r)} \sum_{t=1+r}^{n} [e]^2$$

$$= \frac{1}{(n-r)} e'e \qquad (1.8)$$

et où

$$e = y^{**} - X^{**}\beta; \qquad (1.9)$$

y^{**}: un vecteur colonne de la variable dépendante d'ordre $(n-r) \times 1$,

X^{**}: une matrice de variables indépendantes d'ordre $(n-r) \times j$,

β : un vecteur colonne de paramètres d'ordre $j \times 1$,

e : un vecteur colonne de résidus d'ordre $(n-r) \times 1$.

Si nous substituons (1.8) dans (1.7), nous pouvons poser que:

$$\frac{\partial L(\omega^*)}{\partial \beta} = - \left(\frac{n-r}{2}\right) \frac{1}{\hat{\sigma}^{**2}} \cdot \frac{1}{(n-r)} 2 \left[\frac{\partial e'}{\partial \beta} \cdot e\right] = - \frac{1}{\hat{\sigma}^{**2}} \left[\frac{\partial e'}{\partial \beta} \cdot e\right] \qquad (1.10)$$

$$= \frac{1}{\hat{\sigma}^{**2}} X^{**'} (y^{**} - X^{**}\beta) \qquad (1.11)$$

Et en posant l'expression (1.11) égale à zéro, nous obtenons les es-

timés de β étant donné les paramètres ρ, λ_y, λ_x, λ_e, δ:

$$\hat{\beta}(\rho, \lambda_y, \lambda_x, \lambda_e, \delta) = (X^{**'} X^{**})^{-1} X^{**'} y^{**} \quad . \tag{1.12}$$

Nous obtenons alors les estimateurs β des moindres carrés ordinaires.

Si nous remplaçons β par sa valeur $\hat{\beta}$ dans l'expression (1.9), nous avons pour $\hat{\sigma}^{**2}$ étant donné les paramètres ρ, λ_y, λ_x, λ_e, δ:

$$\hat{\sigma}^{**2}(\rho, \lambda_y, \lambda_x, \lambda_e, \delta) =$$

$$\frac{1}{(n-r)} [y^{**} - X^{**}\hat{\beta}(\rho,\lambda_y,\lambda_x,\lambda_e,\delta)]' [y^{**} - X^{**}\hat{\beta}(\rho,\lambda_y,\lambda_x,\lambda_e,\delta)] \quad . \tag{1.13}$$

<u>Le logarithme de la fonction de vraisemblance concentrée sur σ^{**2} et sur β</u> est obtenu en substituant (1.13) dans (1.6). Ainsi,

$$L(y; \rho, \lambda_y, \lambda_x, \lambda_e, \delta) =$$

$$L(\omega^{**}) = (\lambda_y - 1) \sum_{t=1+r}^{n} \ell n\, y_t - \left(\frac{n-r}{2}\right) \ell n\, \hat{\sigma}^{**2}(\rho,\lambda_y,\lambda_x,\lambda_e,\delta)$$

$$- \frac{1}{2} \sum_{t=1+r}^{n} \ell n\, f(X_t) \quad . \tag{1.14}$$

APPENDICE 2

HYPOTHESE DE NORMALITE DE L'ERREUR.

L'utilisation des transformations paramétriques Box-Cox implique que l'hypothèse de normalité des erreurs résiduelles peut être approximée.

Nous savons que sous l'hypothèse de normalité, le terme d'erreur e_t peut prendre une valeur réelle entre $-\infty$ et $+\infty$. Par conséquent, la variable dépendante peut prendre une valeur dans cette même région. Dans le non-linéaire, la région admissible est plus restrictive.

La transformation paramétrique sur la variable dépendante ne permet pas à celle-ci de prendre n'importe quelle valeur. Comme le note Ramsey (1971), puisque la puissance de transformation ne permet pas au terme d'erreur e_t d'avoir de grandes valeurs réelles négatives, alors la distribution sur e_t ne peut être strictement normale. Les grandes valeurs réelles négatives pour e_t sont donc prohibées. Cependant si la probabilité d'obtenir des valeurs dans la région non permise est minime, nous pouvons noter que le terme d'erreur est approximativement normal.

Nous allons donc montrer que le terme d'erreur e_t suit une distribution normale tronquée lorsque le paramètre de puissance λ_y est supérieur ou inférieur à zéro et suit une distribution normale standard lorsque λ_y tend vers zéro (Tran Cong Liem, 1979, pp.12-14).

Posons à nouveau

$$y_t^{**} = \Sigma_k \beta_k X_{k_t}^{**} + e_t \qquad (2.1)$$

où

$$y_t^{**} = y_t^{*} - \Sigma_\ell \rho_\ell \, y_{t-\ell}^{*} \qquad (2.2)$$

tel que déjà défini dans (40).

Nous pouvons déterminer une condition pour e_t, en termes de la variable dépendante originale y_t, qui permet à la transformation Box-Cox de rester valide (valide si $y_t > 0$).

Si nous substituons (2.2) dans (2.1), l'équation (2.1) devient:

$$y_t^* = \sum_\ell \rho_\ell \, y_{t-\ell}^* + \sum_k \beta_k X_{k_t}^{**} + e_t \qquad (2.3)$$

où

$$y_t^* = \frac{y_t^{(\lambda_y)}}{\sqrt{f(X_t)}}$$

tel que déjà défini dans (42).

Alors,

$$y_t^{(\lambda_y)} = \sqrt{f(X_t)} \left(\sum_\ell \rho_\ell \, y_{t-\ell}^* + \sum_k \beta_k X_{k_t}^{**} + e_t \right) \qquad (2.4)$$

$$y_t = \left[1 + \lambda_y \sqrt{f(X_t)} \left(\sum_\ell \rho_\ell \, y_{t-\ell}^* + \sum_k \beta_k X_{k_t}^{**} + e_t \right) \right]^{(1/\lambda_y)} . \qquad (2.5)$$

Puisque la transformation Box-Cox n'est permise que pour les variables positives, alors nous sommes assurés que la relation (2.5) est strictement positive. Aussi nous avons pour le côté droit de cette dernière expression:

$$\left[1 + \lambda_y \sqrt{f(X_t)} \left(\sum_\ell \rho_\ell \, y_{t-\ell}^* + \sum_k \beta_k X_{k_t}^{**} + e_t \right) \right] > 0 \quad . \qquad (2.6)$$

A partir de (2.5), nous pouvons dériver la condition pour e_t qui nous permet d'avoir $y_t > 0$.

Pour $\lambda_y > 0$, la transformation Box-Cox reste valide si

$$e_t > - \frac{1}{\lambda_y \sqrt{f(X_t)}} - \sum_\ell \rho_\ell \, y_{t-\ell}^* - \sum_k \beta_k X_{k_t}^{**} = e_t^* \qquad (2.7)$$

et

pour $\lambda_y < 0$, si

$$e_t < -\frac{1}{\lambda_y \sqrt{f(X_t)}} - \sum_{\ell} \rho_\ell \, y_{t-\ell}^{*} - \sum_{k} \beta_k X_{k_t}^{**} = e_t^{*} \quad . \tag{2.8}$$

Si nous prenons les valeurs anticipées pour chacun des termes de l'expression (2.1), nous avons:

$$E(y_t^{**}) = \sum_{k} \hat{\beta}_k X_{k_t}^{**} + E(e_t)$$

$$= \hat{y}_t^{**} + E(e_t) \quad , \tag{2.9}$$

où le terme d'erreur e_t suit une distribution normale tronquée dans la région $[e_t^{*}, \infty]$ si $\lambda_y > 0$ et dans la région $[-\infty, e_t^{*}]$ si $\lambda_y < 0$.

Dans la région $[e_t^{*}, \infty]$ avec moyenne:

$$E(e_t) = \hat{e}_t^{*} \int_{-\infty}^{\hat{e}_t^{*}} f_n\,(0, \hat{\sigma}^{**2})\, dz + \int_{\hat{e}_t^{*}}^{+\infty} z\, f_n\,(0, \hat{\sigma}^{**2})\, dz \quad . \tag{2.10}$$

Dans la région $[-\infty, e_t^{*}]$ avec moyenne:

$$E(e_t) = \int_{-\infty}^{\hat{e}_t^{*}} z\, f_n\,(0, \hat{\sigma}^{**2})\, dz + \hat{e}_t^{*} \int_{\hat{e}_t^{*}}^{+\infty} f_n\,(0, \hat{\sigma}^{**2})\, dz \quad . \tag{2.11}$$

où $f_n\,(0, \hat{\sigma}^{**2})$ est la fonction de densité normale univariée avec une moyenne zéro et une variance $\hat{\sigma}^{**2}$,

et

où $\bar{e}_t^{*}$ est la valeur de e_t^{*}, calculée avec les estimateurs (β, ρ, λ_y, λ_x, λ_e, δ) du maximum de vraisemblance.

Si $\lambda_y \to 0$, alors l'expression (2.6) devient par la condition (11')

égale à exp $[\sqrt{f(X_t)}\ (\Sigma_{\ell} \rho_{\ell}\ y_{t-\ell}^{*} + \Sigma_{k} \beta_k X_{k_t}^{**} + e_t)]$. Nous obtenons donc la valeur limite de exp [ ] laquelle est toujours positive. Alors y_t dans l'expression (2.5) sera toujours positif pour $-\infty < e_t < +\infty$ et $E(e_t) = 0$.

Ainsi dans la région $[-\infty, +\infty]$ avec moyenne:

$$E(e_t) = 0 \quad \text{si } \lambda_y \to 0 \quad , \tag{2.12}$$

nous avons une distribution normale non-tronquée.

APPENDICE 3

3 MESURES DE COEFFICIENTS DE CORRELATION MULTIPLE.

. R^2 en termes de y_t et de $E(y_t)$.

La première mesure de R^2 relie les erreurs estimées à la valeur non transformée de la variable dépendante. Ce R^2 donne une mesure intuitive de la qualité de l'ajustement du modèle théorique à la réalité statistique.

$$R^2\,(E(y)) = 1 - \frac{\sum_{t=1+r}^{n} [y_t - E(y_t)]^2}{\sum_{t=1+r}^{n} (y_t - \bar{y})^2} \quad . \tag{3.1}$$

Si nous prenons les valeurs anticipées de l'expression (2.5) de l'appendice 2, nous avons:

$$E(y_t) = [1 + \hat{\lambda}_y \sqrt{f(X_t)} \; (\sum_{\ell} \hat{\rho}_\ell \, \hat{y}_{t-\ell}{}^{*} + \sum_{k} \hat{\beta}_k \hat{X}_{k_t}{}^{**} + E(e_t))]^{1/\hat{\lambda}_y} \quad . \tag{3.2}$$

Le symbole (-) signifie que λ_y, ρ, β et tous les paramètres impliqués dans $f(X_t)$, $y_{t-\ell}{}^{*}$ ainsi que $X_{k_t}{}^{**}$ ont été remplacés par leurs estimateurs du maximum de vraisemblance.

Les deux derniers R^2 relient les erreurs estimées à la valeur transformée de la variable dépendante.

. R^2 en termes de y^{**} et de $E(y^{**})$.

$$R^2\,(E(y^{**})) = 1 - \frac{\sum_{t=1+r}^{n} [y_t{}^{**} - E(y_t{}^{**})]^2}{\sum_{t=1+r}^{n} (y_t{}^{**} - \bar{y}^{**})^2} \quad . \tag{3.3}$$

. R^2 en termes de y^{**} et de $\hat{y}^{**} = X^{**}\hat{\beta}$.

$$R^2(\hat{y}^{**}) = 1 - \frac{\sum_{t=1+r}^{n} (y_t^{**} - \hat{y}_t^{**})^2}{\sum_{t=1+r}^{n} (y_t^{**} - \bar{y}^{**})^2} \quad . \tag{3.4}$$

Les 3 R^2 sont égaux lorsque $\lambda_y = \lambda_{xk} = 1$ en l'absence de l'hétéroscédasticité et de l'autocorrélation.

Ces 3 mesures de coefficients de corrélation multiple sont présentées dans le texte de Tran Cong Liem, (pp.12-14).

APPENDICE 4

DEFINITION DES VARIABLES.

A - VARIABLE DEPENDANTE

VNPEM : consommation d'essence à moteur (en gallons) - ventes nettes mensuelles d'essence.

B - VARIABLES DE PRIX

1. FADCB : tarif réel pour les usagers adultes du mode de transport en commun offert par la C.T.C.U.M..
2. PGRMD : coût réel de l'essence.
3. NCQR : nombre total d'automobiles dans la province de Québec.
4. PCB : indice des prix à la consommation dans la province de Québec.

C - VARIABLE DE REVENU

5. RAWEQ : salaire hebdomadaire réel moyen dans la province de Québec.

D - VARIABLES DE TEMPS

6. TW : temps moyen d'attente en un point aléatoire du réseau pendant une journée moyenne d'un mois donné.
7. TT : temps moyen de transport par le mode de transport en commun en supposant un déplacement représentatif.
8. TAUTO : temps moyen de transport par automobile d'un déplacement représentatif hors pointe sur une distance de quatre milles.

E - VARIABLES DE CONFORT

9. TLK : valeur absolue de la différence entre la température mensuelle moyenne subie et la température représentative d'un mois de septembre particulier (56 degrés Fahrenheit).
10. RFK : précipitation mensuelle de pluie.
11. SFK : précipitation mensuelle de neige.

12. CSFMR : accumulation mensuelle de neige au sol.

F - VARIABLES D'ACTIVITE

13. RCUTPN : inscriptions scolaires totales aux niveaux collégial et universitaire dans la province de Québec.

14. ISECON : inscriptions scolaires totales aux niveaux primaire et secondaire dans la province de Québec.

15. RIDCB : indice réel des ventes au détail.

16. EIQP : indice de l'emploi dans la province de Québec.

17. EX7 : nombre de passagers qui sont entrés à la station de métro située sur le site de l'Exposition universelle de 1967.

18. STR : variable auxiliaire qui mesure l'effet des grèves des employés de la C.T.C.U.M..

19. VWM : nombre de jours de congé d'un travailleur.

G - VARIABLES D'AGREGATION

20. WD : nombre mensuel de jours de travail.

21. SAT : nombre mensuel de samedis.

22. SHD : nombre mensuel de dimanches et de jours fériés.

23. C : constante.

BIBLIOGRAPHIE

1.- Anscombe, F.J. and Tukey, J.W., The Criticism of Transformations. Texte présenté devant the American Statistical Association and Biometric Society, 1954, Montreal.

2.- Beach, C.M. and MacKinnon, J.G., "A Maximum Likelihood Procedure for Regression with Autocorrelated Errors", Econometrica, Vol. 46, 1978, nu. 1, pp.51-58.

3.- Box, G.E.P. and Cox, D.R., "An Analysis of Transformations", Journal of the Royal Statistical Society Series B, Vol. 26, 1964, nu. 2, pp. 211-243.

4.- Box, G.E.P. and Jenkins, G.M., Time Series Analysis Forecasting and Control, Holden-Day, San Francisco, 1970.

5.- Box, G.E.P. and Tildwell, P.W., "Transformation of the Independent Variables", Technometrics, Vol. 4, 1962, nu. 4, pp. 531-550.

6.- Dagum, C. et Dagum, E.B., Construction de modèles et analyse économétrique. Cahiers de l'Institut de sciences mathématiques et économiques appliquées 8, 1974, nos. 11-12.

7.- Dhrymes, P.J., Howrey, E.P. et al., "Criteria for Evaluation of Econometric Models", Annals Econ. Society Measurement, Vol. 1, 1972, pp. 291-324.

8.- Fletcher, R. and Powell, M.J.D., "A Rapidly Convergent Descent Method for Minimization", The Computer Journal, Vol. 6, 1963, pp.163-168.

9.- Gaudry, M.J.I., A Study of Aggregate Bi-Modal Urban Travel Supply, Demand and Network Behavior Using Simultaneous Equations with Autoregressive Residuals, Publication # 88, Centre de recherche sur les transports, Université de Montréal, novembre 1977, révisé en août 1978.

10.- Gaudry, M.J.I. and Dagenais, M.G., The Use of Box-Cox Transformations in Regression Models with Heteroskedastic Autoregressive Residuals, Publication # 106, Centre de recherche sur les transports, Université de Montréal, août 1978, révisé en avril 1979.

11.- Gaudry, M.J.I. and Wills, M.J., Estimating the Functional Form of Travel Demand Models, Publication # 63, Centre de recherche sur les transports, Université de Montréal, juin 1977, révisé en janvier 1978.

12.- Glejser, H., "A New Test for Heteroskedasticity", Journal of the American Statistical Association, Vol. 64, 1969, nu. 325, pp.316-323.

13.- Goldfeld, S.M. and Quandt, R.E., "Some Tests for Homoskedasticity", Journal of the American Statistical Association, Vol. 60, 1965, nu. 310, pp. 539-547.

14.- Goldfeld, S.M. and Quandt, R.E., Nonlinear Methods in Econometrics, ed.

D.W. Jorgenson and J. Walbroeck, Amsterdam, London, North-Holland, 1972.

15.- Heckman, J. and Polachek, S., "Empirical Evidence on The Functional Form of the Earnings-Schooling Relationship", Journal of the American Statistical Association, Vol. 69, 1974, nu. 346, pp.350-354.

16.- Johnston, J., Econometric Methods, McGraw-Hill, 2nd. ed., New-York, 1972, (c) 1963, 437 p..

17.- Kadiyala, K.R., "A Transformation Used to Circumvent the Problem of Autocorrelation", Econometrica, Vol. 36, 1968, nu. 1, pp.93-96.

18.- Kau, J.B. and Sirmans, C.F., The Functional Form of the Gravity Model: A New Technique with Empirical Results. Working Paper. Department of Finance, University of Illinois. Presented at the North American Meeting of the Regional Science Association in Toronto, 1976.

19.- Kendall, M.G. and Stuart, A., The Advanced Theory of Statistics, Vol. II, Griffin, London, 1967.

20.- Major, F., L'hétéroscédasticité dans les modèles agrégés de demande de transport en commun en séries chronologiques, Publication # 102, Centre de recherche sur les transports, Université de Montréal, juin 1978.

21.- Park, R.E., "Estimation with Heteroskedastic Error Terms", Econometrica, Vol. 34, 1966, nu. 4, p.888.

22.- Ramsey, J.B. and Zarembka, P., "Specification Error Tests and Alternative Functional Forms of the Aggregate Production Function", Journal of the American Statistical Association, Vol. 66, pp.471-477.

23.- Savin, N.E. and White, K.J., "Estimation and Testing for Functional Form and Autocorrelation: A Simultaneous Approach", Journal of Econometrics, Vol. 8, 1978, pp.1-12.

24.- Schlesselmann, J., "Power Families: A Note on the Box and Cox Transformation", Journal of the Royal Statistical Society Series B, Vol. 33, 1971, nu. 2, pp.307-311.

25.- Spitzer, J.J., "The Demand for Money, the Liquidity Trap, and Functional Forms", International Economic Review, Vol. 17, 1976, pp.220-227.

26.- Spitzer, J.J., "A Simultaneous Equations System of Money Demand and Supply Using Generalized Functional Forms", Journal of Econometrics, Vol. 5, 1977, pp.117-128.

27.- Theil, H., "Estimates and their Sampling Variance of Parameters of Certain Heteroskedastic Distributions", dans Revue de l'Institut International de Statistique, Vol. 19, 1951, pp.141-147.

28.- Tran Cong, Liem, DSDM: L-1, A Program for Box-Cox Transformations in Regression Models with Heteroskedastic and Autoregressive Residuals, Publication # 134, Centre de recherche sur les transports, Université de Montréal, juin 1979.

UNIVERSITÉ DE MONTRÉAL

29.- Tukey, J.W., "On the Comparative Anatomy of Transformations", Annals of Mathematical Statistics, Vol. 28, 1957, pp.602-632.

30.- Turner, M.E. et al., "Generalized Asymptotic Regression and Nonlinear Path Analysis", Biometrics, Vol. 17, 1961, pp.120-143.

31.- Welland, J.D., Cognitive Abilities, Schooling and Earnings: The Question of Functional Form. Working Paper 76-14. Department of Economics, McMaster University, Hamilton, 1976.

32.- Zarembka, P., "Functional Form in the Demand for Money", Journal of the American Statistical Association, Vol. 63, 1968, nu. 322, pp.502-511.

33.- Zarembka, P., Transformation of Variables in Econometrics, Ch. 3 of Frontiers in Econometrics, ed. P. Zarembka, 1974, pp.81-104.

UNIVERSITÉ DE MONTRÉAL

REMERCIEMENTS

Je tiens à témoigner ma gratitude envers le Professeur Marc Gaudry de sa direction constante et éclairée dans l'élaboration de ce rapport de recherche.

Je veux également remercier le Professeur Marcel Dagenais de ses conseils judicieux et de sa participation à la correction de ce rapport.

Je tiens aussi à exprimer ma reconnaissance à M. Tran Cong Liem, assistant de recherche, pour son assistance technique.